ARCHIVE
for
RATIONAL MECHANICS
and
ANALYSIS

Edited by

C. TRUESDELL

Volume 5, Number 3

SPRINGER-VERLAG
BERLIN-GÖTTINGEN-HEIDELBERG
(Postverlagsort Berlin · 17. 5. 1960)

Mechanicam vero duplicem Veteres constituerunt: Rationalem *quae per Demonstrationes accurate procedit, & Practicam. Ad practicam spectant Artes omnes Manuales, a quibus utique* Mechanica *nomen mutuata est. Cum autem Artifices parum accurate operari soleant, fit ut* Mechanica *omnis a* Geometria *ita distinguatur, ut quicquid accuratum sit ad* Geometriam *referatur, quicquid minus accuratum ad* Mechanicam. *Attamen errores non sunt Artis sed Artificum. Qui minus accurate operatur, imperfectior est Mechanicus, & si quis accuratissime operari posset, hic foret Mechanicus omnium perfectissimus.* NEWTON

La généralité que j'embrasse, au lieu d'éblouïr nos lumieres, nous découvrira plutôt les véritables loix de la Nature dans tout leur éclat, & on y trouvera des raisons encore plus fortes, d'en admirer la beauté & la simplicité. EULER

Ceux qui aiment l'Analyse, verront avec plaisir la Méchanique en devenir une nouvelle branche ... LAGRANGE

The ARCHIVE FOR RATIONAL MECHANICS AND ANALYSIS nourishes the discipline of mechanics as a deductive, mathematical science in the classical tradition and promotes pure analysis, particularly in contexts of application. Its purpose is to give rapid and full publication to researches of exceptional moment, depth, and permanence.

Each memoir must meet a standard of rigor set by the best work in its field Contributions must consist largely in original research; on occasion, an expository paper may be invited.

English, French, German, Italian, and Latin are the languages of the Archive. Authors are urged to write clearly and well, avoiding an excessively condensed or crabbed style.

Manuscripts intended for the Archive should be submitted to an appropriate member of the Editorial Board.

The ARCHIVE FOR RATIONAL MECHANICS AND ANALYSIS appears in numbers struck off as the material reaches the press; five numbers constitute a volume. Subscriptions may be entered through any agent. The price is DM 96.— per volume.

For each paper 75 offprints are provided free of charge.

A New Proof of the Bieberbach Conjecture for the Fourth Coefficient

Z. CHARZYNSKI & M. SCHIFFER

Introduction

In 1955 GARABEDIAN & SCHIFFER [1] proved that the fourth coefficient in the Taylor series of a function $f(z)$, univalent in the unit circle, satisfies the inequality $|a_4| \leqq 4$. The argument consisted in two basic assertions: a) It was shown by means of the calculus of variations and by use of LOEWNER's differential equation that the problem of finding an extremum function for the Bieberbach problem is equivalent to a nonlinear boundary-value problem for a system of ordinary differential equations. It was proved also that the Koebe function led to a solution of this boundary value problem and that no other solution could lie in a very small neighborhood of this particular solution. b) There remained then the task of showing that outside of this critical neighborhood there could exist no solution leading to an extremum function of the Bieberbach problem. This was achieved by a laborious and complicated set of estimates and inequalities.

While the approach of that paper shows in principle a general method of attack upon coefficient problems for schlicht functions, the amount of labor involved in this relatively simple question seemed discouraging, and it was doubtful whether higher coefficients should be dealt with in a corresponding manner. The purpose of the present paper is to give a great simplification of the proof that $|a_4| \leqq 4$ and to show at the same time that more elementary methods are sufficient to establish the theorem. We start from an inequality implicit in a paper of GRUNSKY [2] of 1939. This inequality can be derived by elementary means, using the Schwarz inequality and standard methods of contour integration. Below we give a new proof of the Grunsky inequality based on variational methods, both in order to make this paper self-consistent, and also since the new proof seems somewhat shorter than GRUNSKY's. No use of LOEWNER's differential equation is necessary. We hope that our present proof for the Bieberbach conjecture for a_4 will encourage the attack upon higher coefficients and stimulate new interest in this beautiful field of function theory.

§ 1. The fundamental inequalities

1. We shall be dealing with extremum problems within the family $\mathfrak{S}$ of all functions $f(z)$ which are regular and univalent for $|z| < 1$ and which possess a Taylor series

$$f(z) = z + a_2 z^2 + a_3 z^3 + \cdots + a_n z^n + \cdots. \tag{1}$$

Let $w=f(z)$ yield the image domain D of the unit circle, and let w_0 be an arbitrary point in the w-plane but not in D. There exist for arbitrarily small values of $\varrho>0$ infinitely many members of $\mathfrak{S}$ and of the form [3]

$$f^*(z)=f(z)+\frac{a\,\varrho^2 f(z)^2}{(f(z)-w_0)\,w_0}+O(\varrho^3), \qquad |a|<1, \tag{2}$$

where the correction term $O(\varrho^3)$ can be estimated uniformly in each closed subdomain $|z|\leq r<1$. These functions $f^*(z)$ may then serve as comparison functions with the original member $f(z)$ in variational problems regarding the family $\mathfrak{S}$.

We shall be interested in maximum problems concerning the first three coefficients a_2, a_3, and a_4 in (1). Therefore it is necessary to express the change of these coefficients under the variation (2). An easy calculation yields

$$\begin{aligned} a_2^* &= a_2-\frac{a\,\varrho^2}{w_0^3}+O(\varrho^3)\\ a_3^* &= a_3-\frac{a\,\varrho^2}{w_0^3}\left(2a_2+\frac{1}{w_0}\right)+O(\varrho^3)\\ a_4^* &= a_4-\frac{a\,\varrho^2}{w_0^3}\left(2a_3+a_2^2+\frac{3a_2}{w_0}+\frac{1}{w_0^2}\right)+O(\varrho^3). \end{aligned} \tag{3}$$

Now let $F=F(a_2, a_3, a_4)$ be an analytic function of the three complex variables a_2, a_3, and a_4. F is then a functional of $f\in\mathfrak{S}$ with the corresponding variational formula

$$\begin{aligned} F^* = F &- \frac{a\,\varrho^2}{w_0^3}\left[\frac{\partial F}{\partial a_4}(2a_3+a_2^2)+\frac{\partial F}{\partial a_3}2a_2+\frac{\partial F}{\partial a_2}+\right.\\ &\left.+\left(\frac{\partial F}{\partial a_4}3a_2+\frac{\partial F}{\partial a_3}\right)\frac{1}{w_0}+\frac{\partial F}{\partial a_4}\cdot\frac{1}{w_0^2}\right]+O(\varrho^3). \end{aligned} \tag{4}$$

Having chosen a fixed functional $F(a_2, a_3, a_4)$, we may ask for the solution of the extremum problem

$$\mathrm{Re}\{F\}=\max \qquad \text{within the family } \mathfrak{S}. \tag{5}$$

The existence of extremizing functions in $\mathfrak{S}$ follows from the compactness of the family. We can characterize the extremum functions $f(z)$ by the fact that under all variations (2) the inequality $\mathrm{Re}\{F^*\}\leq \mathrm{Re}\{F\}$ must hold. In other words,

$$\begin{aligned} &\mathrm{Re}\left\{\frac{a\,\varrho^2}{w_0^3}\left[A+\frac{B}{w_0}+\frac{C}{w_0^2}\right]\right\}\geq O(\varrho^3)\\ &A=(2a_3+a_2^2)\frac{\partial F}{\partial a_4}+2a_2\frac{\partial F}{\partial a_3}+\frac{\partial F}{\partial a_2},\\ &B=3a_2\frac{\partial F}{\partial a_4}+\frac{\partial F}{\partial a_3}, \qquad C=\frac{\partial F}{\partial a_4} \end{aligned} \tag{6}$$

must hold for every admissible variation (2).

The basic lemma of the calculus of variations within the class $\mathfrak{S}$ leads now to the conclusion that D is a slit domain bounded by analytic arcs Γ with a parametric representation $w=w(\tau)$ satisfying the differential equation

$$\left(\frac{dw}{d\tau}\right)^2\cdot\frac{1}{w^3}\left[A+\frac{B}{w}+\frac{C}{w^2}\right]<0. \tag{7}$$

So far, every extremum problem (5) can be brought within the characterization (7). It is, however, extremely difficult to draw decisive conclusions from (7) and to determine precisely the extremum function sought. Indeed, the coefficient problem of univalent functions is contained in this particular problem. Since it has not been possible to find a satisfactory general solution of (7), we may now ask for particular functionals F for which the complete answer can be given. It is clear from the form of (7) that the integration of the differential equation will be simplified if the term in brackets is a perfect square. We ask, therefore, for those functions $F(a_2, a_3, a_4)$ for which

$$(8)\quad B^2 = 4AC, \text{ i.e., } \left(3a^2 \frac{\partial F}{\partial a_4} + \frac{\partial F}{\partial a_3}\right)^2 = 4\frac{\partial F}{\partial a_4}\left[(2a_3 + a_2^2)\frac{\partial F}{\partial a_4} + 2a_2 \frac{\partial F}{\partial a_3} + \frac{\partial F}{\partial a_2}\right].$$

This condition is a simple first order partial differential equation for F. We can give immediately a particular integral of it which is in polynomial form and which contains an arbitrary constant of integration l. Namely,

$$(9)\qquad F(a_2, a_3, a_4) = a_4 - 2a_2 a_3 + \tfrac{13}{12} a_2^3 + 2l(a_3 - \tfrac{3}{4} a_2^2) + l^2 a_2.$$

Conversely, if we start with the functional (9) and ask for the solution of the extremum problem (5), we arrive at the differential equation for the boundary curve Γ:

$$(10)\qquad \left(\frac{dw}{d\tau}\right)^2 \cdot \frac{1}{w^3}\left[l + \frac{1}{2} a_2 + \frac{1}{w}\right]^2 < 0.$$

Let $w = f(e^{i\tau})$ be a particular parametric representation for Γ. Now

$$(11)\qquad \frac{dw}{d\tau} = i\, e^{i\tau} f'(e^{i\tau}).$$

Hence, if we denote

$$(12)\qquad \frac{z^2 f'(z)^2}{f(z)^3}\left[l + \frac{1}{2} a_2 + \frac{1}{f(z)}\right]^2 = q(z),$$

we obtain in $q(z)$ an analytic function of z in the unit circle which is regular there except for a pole of order 3 at the origin and which has non-negative values on the boundary. Hence, by SCHWARZ' reflection principle, $q(z)$ is rational in the entire z-plane and has the form

$$(13)\qquad q(z) = \frac{1}{z^3} + \frac{\alpha}{z^2} + \frac{\beta}{z} + \gamma + \bar{\beta} z + \bar{\alpha} z^2 + z^3, \qquad \gamma \text{ real}.$$

We can easily factorize $q(z)$. It can vanish for $|z| < 1$ only at the point z_0 for which

$$(14)\qquad f(z_0)\,(l + \tfrac{1}{2} a_2) = -1.$$

If we put $z_0 = \varrho\, e^{i\varphi}$, the point $\bar{z}_0^{-1} = \frac{1}{\varrho}\, e^{i\varphi}$ will likewise be a zero of $q(z)$. Obviously, both points are double zeros of $q(z)$. Each root of $q(z)$ on the unit circle must be double, since $q(z)$ does not change sign on this curve. Hence

$$(15)\qquad z^3 q(z) = A^2 (z - e^{i\psi})^2 (z - \varrho\, e^{i\varphi})^2 \left(z - \frac{1}{\varrho}\, e^{i\varphi}\right)^2$$

if a root z_0 of (14) lies in the unit circle, or

$$(15') \qquad z^3 q(z) = A^2 (z - e^{i\varphi})^2 (z - e^{i\varphi_1})^2 (z - e^{i\varphi_2})^2$$

if all roots of $q(z)$ lie on $|z| = 1$.

In any case, in view of (13) we can assert that

$$(16) \qquad \sqrt{q(z)} = \frac{\varkappa z^3 + \mu z^2 + \nu z + 1}{z^{\frac{3}{2}}}$$

and that this expression is real for $|z| = 1$. Hence we may conclude that $\varkappa = 1$, $\mu = \bar{\nu}$, and we may put (12) into the form

$$(17) \qquad \frac{z^{\frac{5}{2}} f'(z)}{f(z)^{\frac{5}{2}}} \left[\left(l + \frac{1}{2} a_2\right) f(z) + 1\right] = z^3 + \mu z^2 + \bar{\mu} z + 1 .$$

We insert into the left-hand side of (17) the series development (1) of $f(z)$ and compute the development of this term at the origin. We find that

$$(18) \qquad \begin{aligned} \frac{z^{\frac{5}{2}}}{f(z)^{\frac{5}{2}}} f'(z) \left[1 + \left(l + \frac{1}{2} a_2\right) f(z)\right] &= 1 + l z + \left(\frac{1}{2} a_3 - \frac{3}{8} a_2^2 + \frac{1}{2} l a_2\right) z^2 + \\ &+ \left[\frac{3}{2} a_4 - 3 a_2 a_3 + \frac{13}{8} a_2^3 + l \left(\frac{3}{2} a_3 - \frac{9}{8} a_2^2\right)\right] z^3 + \cdots . \end{aligned}$$

Comparing coefficients in (17) and (18), we obtain

$$(19) \qquad \begin{gathered} l = \bar{\mu} \\ (a_3 - \tfrac{3}{4} a_2^2) + l a_2 = 2\mu \\ a_4 - 2 a_2 a_3 + \tfrac{13}{12} a_2^3 + l (a_3 - \tfrac{3}{4} a_2^2) = \tfrac{2}{3} . \end{gathered}$$

We multiply the second equality (19) by $l = \bar{\mu}$ and add it to the third equality. We arrive finally at the result

$$(20) \qquad F = a_4 - 2 a_2 a_3 + \tfrac{13}{12} a_2^3 + 2 l (a_3 - \tfrac{3}{4} a_2^2) + l^2 a_2 = \tfrac{2}{3} + 2 |l|^2 .$$

The right-hand side of (20) depends only on the parameter l, which we can choose arbitrarily at the beginning of the inquiry. The left-hand side is the value of the functional in the extremum case. Hence we find the general coefficient inequality, valid for all functions of the family $\mathfrak{S}$:

$$(21) \qquad \operatorname{Re}\{a_4 - 2 a_2 a_3 + \tfrac{13}{12} a_2^3 + 2 l (a_3 - \tfrac{3}{4} a_2^2) + l^2 a_2\} \leqq \tfrac{2}{3} + 2 |l|^2 .$$

This inequality is fundamental for our further applications. It asserts a relation between a quadratic expression on the left and a hermitian expression in l on the right. It thus reminds one of the important inequalities due to Grunsky [2] for coefficients of univalent functions. It can, indeed, be derived from the Grunsky inequalities. This fact is of great methodological interest since Grunsky's inequalities are derived without use of variational methods and are closely related to the classical area theorem. We gave the above direct proof of (21) in order to show the significance of the Grunsky inequalities in the theory of variations, where they are closely related to the problem of differential equations containing perfect squares [4].

2. We shall later need a simple inequality for the coefficients in $\mathfrak{S}$ which follows from the area theorem. Consider

$$g(z) = \frac{1}{\sqrt{f\left(\frac{1}{z^2}\right)}} = z + \frac{b_1}{z} + \frac{b_3}{z^3} + \cdots. \tag{22}$$

Clearly, $g(z)$ is regular and univalent for $|z| > 1$, and an easy computation yields

$$b_1 = -\tfrac{1}{2} a_2, \quad b_3 = -\tfrac{1}{2}\left(a_3 - \tfrac{3}{4} a_2^2\right). \tag{23}$$

We have the area theorem:

$$1 \geqq |b_1|^2 + 3|b_3|^2 + 5|b_5|^2 + \cdots. \tag{24}$$

The combination

$$\lambda = a_3 - \tfrac{3}{4} a_2^2 \tag{25}$$

will occur frequently in our calculations. We derive for it from (23) and (24) the inequality

$$3|\lambda|^2 \leqq 4 - |a_2|^2. \tag{26}$$

§ 2. Proof that $|a_4| \leqq 4$

1. We wish to show now that the inequalities (1.21) and (1.26) imply the Bieberbach inequality for the fourth coefficient. We may assume without loss of generality that

$$a_4 \geqq 0, \quad \operatorname{Re}\{a_2\} \geqq 0, \tag{1}$$

since this case can always be achieved by replacing $f(z)$ by the new member of $\mathfrak{S}$: $e^{-i\alpha} f(e^{i\alpha} z)$.

We set

$$\lambda = \xi + i\eta, \quad a_2 = 2 - x + iy, \quad 0 \leqq x \leqq 2. \tag{2}$$

We write (1.21) with l as a real parameter at our disposal:

$$a_4 \leqq \tfrac{2}{3} + \operatorname{Re}\{2a_2\lambda\} + \tfrac{5}{12}\operatorname{Re}\{a_2^3\} - 2l\operatorname{Re}\{\lambda\} + l^2\left(2 - \operatorname{Re}\{a_2\}\right). \tag{3}$$

We specialiez by selecting l so as to give the minimum value to the right-hand side:

$$l = \frac{\operatorname{Re}\{\lambda\}}{2 - \operatorname{Re}\{a_2\}}. \tag{4}$$

Now (3) becomes

$$a_4 \leqq \frac{2}{3} + \operatorname{Re}\{2a_2\lambda\} + \frac{5}{12}\operatorname{Re}\{a_2^3\} - \frac{\operatorname{Re}\{\lambda\}^2}{2 - \operatorname{Re}\{a_2\}}. \tag{5}$$

We insert here the real and imaginary parts according to the definition (2):

$$a_4 \leqq \frac{2}{3} + 2\{(2-x)\xi - y\eta\} + \frac{5}{12}\{(2-x)^3 - 3(2-x)y^2\} - \frac{\xi^2}{x}. \tag{6}$$

Observe that

$$2(2-x)\xi - \frac{\xi^2}{x} \leqq (2-x)^2 x. \tag{7}$$

Hence

(8) $$a_4 \leqq 4 + (2-x)^2 x - 5x + \tfrac{5}{2}x^2 - \tfrac{5}{12}x^3 - (\tfrac{5}{2} - \tfrac{5}{4}x)\, y^2 - 2y\eta.$$

By using the identity

(9) $$2y\eta = \alpha\left(y + \frac{1}{\alpha}\eta\right)^2 - \alpha y^2 - \frac{1}{\alpha}\eta^2, \qquad \alpha > 0,$$

we may transform (8) into

(10) $$a_4 - 4 \leqq -x - \frac{3}{2}x^2 + \frac{7}{12}x^3 - \left(\frac{5}{2} - \frac{5}{4}x\right) y^2 + \alpha y^2 + \frac{1}{\alpha}\eta^2.$$

We use next the inequality (1.26), from which, expressed in the notation (2), follows the estimate

(11) $$\eta^2 \leqq \tfrac{4}{3}x - \tfrac{1}{3}(x^2 + y^2).$$

Hence, we can eliminate η from (10) and obtain

(12) $$a_4 - 4 \leqq \left(\frac{4}{3\alpha} - 1\right) x - \left(\frac{3}{2} + \frac{1}{3\alpha}\right) x^2 + \frac{7}{12}x^3 - \left(\frac{5}{2} + \frac{1}{3\alpha} - \alpha - \frac{5}{4}x\right) y^2.$$

The final proof of the inequality $a_4 \leqq 4$ is achieved by two estimates. We first choose

(13) $$\alpha = \tfrac{4}{3}$$

and find that

(14) $$a_4 - 4 \leqq -\frac{7}{4}x^2\left(1 - \frac{x}{3}\right) - \left(\frac{17}{12} - \frac{5}{4}x\right) y^2.$$

We can infer from this inequality that

(15) $$a_4 \leqq 4 \quad \text{when} \quad 0 \leqq x \leqq \tfrac{17}{15}.$$

Next, we select $\alpha = \frac{1}{2}$ and obtain

(16) $$a_4 - 4 \leqq \tfrac{5}{3}x - \tfrac{13}{6}x^2 + \tfrac{7}{12}x^3 - (\tfrac{8}{3} - \tfrac{5}{4}x)\, y^2.$$

The coefficient of y^2 is negative for all values $0 \leqq x \leqq 2$; hence, everything depends on the sign of

(17) $$r(x) = 5 - \tfrac{13}{2}x + \tfrac{7}{4}x^2.$$

One verifies directly that

(18) $$r(2) = -1, \qquad r(\tfrac{17}{15}) = -\tfrac{107}{900}.$$

Thus $r(x)$ must be negative in the entire interval $\frac{17}{15} \leqq x \leqq 2$. Hence we have proved that

(19) $$a_4 \leqq 4 \quad \text{when} \quad \tfrac{17}{15} \leqq x \leqq 2.$$

This completes the proof of the Bieberbach inequality

(20) $$|a_4| \leqq 4.$$

2. For the sake of completeness we wish to indicate briefly how the fundamental inequality (1.21) follows from GRUNSKY's inequality. Let

(21) $$g(z) = z + \sum_{\nu=1}^{\infty} \frac{b_\nu}{z^\nu}$$

be univalent in the circular domain $|z|>1$. Define the n^{th} Faber polynomial of $g(z)$ as the polynomial of degree n of $g(z)$ for which we have the series development

$$G_n[g] = z^n + \sum_{\nu=1}^{\infty} \frac{b_{n\nu}}{z^\nu}. \tag{22}$$

It is easily seen that $G_n(t)$ is uniquely determined by this condition. GRUNSKY proved that the set of inequalities between quadratic and hermitian forms,

$$\left|\sum_{n,m=1}^{N} m\, b_{nm}\, x_n\, x_m\right| \leqq \sum_{n=1}^{N} n\,|x_n|^2, \qquad N = 1, 2, \ldots, \tag{23}$$

are necessary and sufficient conditions that the function (21) be univalent.

We can apply this result to the particular function (1.22), choose $x_1=l$, $x_2=0$, $x_3=\frac{1}{3}$, and find precisely the required inequality (1.21).

This work was performed partially under an Office of Naval Research contract, Nonr-225(11), with Stanford University.

References

[1] GARABEDIAN, P. R., & M. SCHIFFER: A proof of the Bieberbach conjecture for the fourth coefficient. J. Rational Mech. Anal. **4**, 427—465 (1955).

[2] GRUNSKY, H.: Koeffizientenbedingungen für schlicht abbildende meromorphe Funktionen. Math. Z. **45**, 29—61 (1939).

[3] SCHIFFER, M.: A method of variation within the family of simple functions. Proc. London Math. Soc. **44**, 432—449 (1938).

[4] SCHIFFER, M.: Faber polynomials in the theory of univalent functions. Bull. Amer. Math. Soc. **54**, 503—517 (1948).

Stanford University
California

(Received March 7, 1960)

Global Asymptotic Stability for Nonlinear Systems of Differential Equations and Applications to Reactor Dynamics

J. J. LEVIN & J. A. NOHEL

Communicated by L. CESARI

1. Introduction

In this paper[1] conditions are obtained under which all solutions of certain real nonlinear systems of differential equations tend to zero as $t \to \infty$. This study originated in some problems of reactor dynamics; however, the systems investigated are of a quite general nature. Therefore, the results are given in an abstract setting and only in the last section interpreted physically.

Theorem 1 is concerned with the following general form of the Liénard equation:

(1.1) $$\ddot{x} + h(t, x, \dot{x})\,\dot{x} + f(x) = e(t) \qquad (\cdot = d/dt).$$

Theorem 1. *Let the following conditions be satisfied:*

(1.2) $h(t, x, z)$, $f(x)$, $e(t)$ *are sufficiently smooth for a local existence and uniqueness theorem to hold for* (1.1) *on* $0 \leqq t < \infty$; $-\infty < x,\ z < \infty$.

(1.3) *There exists a constant* $k > 0$ *such that*

$$k \leqq h(t, x, z) \qquad (0 \leqq t < \infty;\ -\infty < x,\ z < \infty).$$

(1.4) *Given any constant* $B > 0$ *there exists a constant* $K_B > 0$ *such that*

$$h(t, x, z) \leqq K_B \qquad (0 \leqq t < \infty;\ |x|, |z| \leqq B).$$

(1.5) *Given any constant* $B > 0$ *there exists a constant* $K_B > 0$ *such that*

$$|f(x_1) - f(x_2)| \leqq K_B |x_1 - x_2| \qquad (|x_1|, |x_2| \leqq B).$$

(1.6) $$x f(x) > 0 \qquad (x \neq 0).$$

(1.7) $$g(x) = \int_0^x f(\xi)\, d\xi \to \infty \quad \text{as} \quad |x| \to \infty.$$

---

[1] Supported by Lincoln Laboratory, a center for research operated by Massachusetts Institute of Technology with the joint support of the U.S.Army, Navy, and Air Force, by a grant from the National Science Foundation, and by the Oak Ridge National Laboratory.

(1.8) *There exists a constant* $K > 0$ *such that*

$$|e(t)| \leqq K \qquad (0 \leqq t < \infty).$$

(1.9)
$$\int_0^\infty |e(t)|\, dt < \infty.$$

Then given any x_0, $\dot{x}_0$ *the solution* $x(t)$ *of* (1.1) *satisfying* $x(0) = x_0$, $\dot{x}(0) = \dot{x}_0$ *exists on* $0 \leqq t < \infty$ *and*

(1.10)
$$\lim_{t\to\infty} x(t) = 0, \qquad \lim_{t\to\infty} \dot{x}(t) = 0.$$

An immediate consequence of Theorem 1 is

Corollary 1. *If in addition to the hypothesis of Theorem* 1 *the condition* $e(t) \to 0$ *as* $t \to \infty$ *is also satisfied, then* $\ddot{x}(t) \to 0$ *as* $t \to \infty$.

As will be seen from the proofs, the uniqueness hypothesis (in all the theorems) is not really necessary and is used only to simplify various statements.

If $e(t) \equiv 0$, then Theorem 1 may be thought of as a "global" asymptotic stability theorem of the trivial solution $x(t) = 0$. That is, all solutions tend to zero as $t \to \infty$ and not merely those for which $|x_0|$, $|\dot{x}_0|$ are sufficiently small. The presence of the $e(t)$ term causes, of course, $x(t) \equiv 0$ no longer to be a solution of (1.1). This sort of complication of stability problems has been considered, for example, in [*1*, ch. 13].

The proof of Theorem 1 is subdivided into three parts, Lemmas 2.1, 2.2, and 2.3. These establish respectively (under an increasing number of hypotheses): the existence and boundedness (the bound depending on x_0, $\dot{x}_0$) of $x(t)$, $\dot{x}(t)$ on $0 \leqq t < \infty$; $\dot{x}(t) \to 0$ as $t \to \infty$; and finally $x(t) \to 0$ as $t \to \infty$. Lemma 2.1, which is of interest in itself, is due to ANTOSIEWICZ [*2*] in the special case that h is independent of t. We give a similar proof of this lemma, since the method and some of the formulae are basic to all that follows. The method involves the same sort of considerations as in the Liapounov second method. However, as in [*2*], one cannot simply cite the classical Liapounov ordinary stability theorem, see [*3*, p. 113], in order to obtain Lemma 2.1, for a global rather than a local result is desired. The situation is further complicated here in that the total derivative of the energy function from which (2.6) is obviously obtained is not negative definite (even for the case $e(t) \equiv 0$, see (2.7)). Thus the classical Liapounov asymptotic stability theorem, see [*3*, p. 114], cannot be applied here to give even a local asymptotic result. Lemmas 2.2 and 2.3 are used to circumvent these difficulties.

As part of the novelty of Theorem 1 lies in the fact that h is allowed to depend on t, it is of interest to focus some attention on hypotheses (1.3) and (1.4) which relate to this dependence. That (1.3) cannot be dropped entirely is obvious from the example $\ddot{x} + x = 0$. It is not quite so transparent that (1.4) cannot be dispensed with, especially as (1.4) is automatically satisfied if h is continuous and does not depend on t. Remarks 2.1 and 2.2, which deal with a linear case of (1.1) and are in the spirit of the present methods, exhibit interesting examples in which (1.3) and (1.4) are respectively the only hypotheses of Theorem 1 that are violated; in both, the conclusion of Theorem 1 does not hold. In connection with Remark 2.2 see [*4*, p. 137].

Theorem 2 is concerned with the system

$$\begin{aligned} \dot{x} &= -\sum_{i=1}^{n} a_i z_i, \\ \dot{z}_i &= -h_i(t, x, z)\, z_i + b_i f(x) + e_i(t) \qquad (i = 1, \dots, n), \end{aligned} \tag{1.11}$$

where the a_i and b_i are constants and $z = (z_1, \dots, z_n)$, which for $n = 1$ is essentially (1.1).

Theorem 2. *Let the functions h_i, f, e_i be sufficiently smooth for a local existence and uniqueness theorem to hold for* (1.11) *on $0 \leqq t < \infty$; $-\infty < x, z_i < \infty$. Let the h_i satisfy* (1.3, 1.4), *f* (1.5, 1.6, 1.7), *and the e_i* (1.8, 1.9). *Furthermore, let the constants a_i, b_i satisfy either*

(1.12) $a_i = c\, b_i$, *where $c > 0$ and the b_i are arbitrary except that at least one, say b_j, is not zero,*

or

$$a_i / b_i > 0 \qquad (i = 1, \dots, n). \tag{1.13}$$

Then given any x_0, z_0, the solution $x(t)$, $z(t)$ of (1.11) *satisfying $x(0) = x_0$, $z(0) = z_0$ exists on $0 \leqq t < \infty$, and*

$$\lim_{t \to \infty} x(t) = 0, \qquad \lim_{t \to \infty} z(t) = 0. \tag{1.14}$$

In (1.3) $-\infty < z < \infty$ means $-\infty < z_i < \infty$ $(i = 1, \dots, n)$, and $|z|$ in (1.4) means $|z| = \sum |z_i|$.

The proof of Theorem 2 differs only in minor details from that of Theorem 1. Therefore, analogues of Lemmas 2.1, 2.2, 2.3 are not stated explicity and only the essential formulae involved in the proof are given. An obvious analogue of Corollary 1 holds here when $e_i(t) \to 0$ as $t \to \infty$.

Consideration of the linear special case of (1.11) defined by

$$h_i(t, x, z) \equiv h_i > 0, \qquad f(x) = x, \qquad e_i(t) \equiv 0, \tag{1.15}$$

where the h_i are constants, sheds additional light on conditions (1.12, 1.13). It is shown, Lemma 3.1, that the real part of each of the characteristic roots of the coefficient matrix associated with (1.11, 1.15) is negative if either (1.12) or (1.13) is satisfied. Hence, in the special case of (1.15), Theorem 2 reduces to a well-known theorem of LIAPOUNOV, see [*1*, p. 314]. This classical theorem may be used together with a standard perturbation technique to obtain a local (but not global) asymptotic stability result if, for example, the second condition in (1.15) is generalized to $f(x) = x + o(x)$ as $x \to 0$. However, a perturbation technique is hopeless if, say, $f(x) = x^3$ (which satisfies (1.5, 1.6, 1.7)).

We remark that the last n equations of (1.11) may be written as

$$\dot{z} = G(t, x, z)\, z + b\, f(x) + e(t),$$

where

$$G = -\operatorname{diag}(h_1, \dots, h_n), \quad z = \operatorname{col}(z_1, \dots, z_n), \quad b = \operatorname{col}(b_1, \dots, b_n), \quad e = \operatorname{col}(e_1, \dots, e_n).$$

However, if G is of the form $G = -h(t, x, z)\, A$, where $h(t, x, z)$ is a scalar function satisfying (1.3, 1.4) and A is a positive definite real symmetric matrix, then

the system is easily transformed into a special case of (1.11). A similar comment applies to (1.16) below. This observation is used in the reactor applications of Section 5.

Theorem 3 is concerned with the system

$$
\begin{aligned}
\dot{x} &= -a_0 y - \sum_{i=1}^{n} a_i z_i, \\
\dot{y} &= b_0 f(x) + e_0(t), \qquad\qquad (1.16) \\
\dot{z}_i &= -h_i(t, x, y, z) z_i + b_i f(x) + e_i(t) \qquad (i = 1, \ldots, n),
\end{aligned}
$$

where again the a_i and b_i are constants and $z = (z_1, \ldots, z_n)$.

Theorem 3. *Let the functions h_i, f, e_i be sufficiently smooth for a local existence and uniqueness theorem to hold for* (1.16) *on $0 \leqq t < \infty$; $-\infty < x, y, z_i < \infty$. Let the h_i satisfy* (1.3, 1.4), *f* (1.5, 1.6, 1.7), *and the e_i* (1.8, 1.9). *Furthermore, let the constants a_i, b_i satisfy either*

(1.17) *$a_i = c\, b_i$, where $c > 0$ and the b_i are arbitrary except that $b_0 \neq 0$ and at least one other b_i, say b_j, is not zero,*

or

$$a_i / b_i > 0 \qquad (i = 0, \ldots, n). \qquad (1.18)$$

Then given any x_0, y_0, z_0, the solution $x(t)$, $y(t)$, $z(t)$ of (1.16) *satisfying $x(0) = x_0$, $y(0) = y_0$, $z(0) = z_0$ exists on $0 \leqq t < \infty$, and*

$$\lim_{t\to\infty} x(t) = 0, \qquad \lim_{t\to\infty} y(t) = 0, \qquad \lim_{t\to\infty} z(t) = 0. \qquad (1.19)$$

Here the range of the variables in (1.3) is $0 \leqq t < \infty$; $-\infty < x, y, z_i < \infty$, and the condition in (1.4) is $0 \leqq t < \infty$; $|x|, |y|, |z| \leqq B$.

The first part of the proof of Theorem 3 again follows Lemmas 2.1 and 2.2 in all essentials. However, the last part of the proof, while having some similarity to Lemma 2.3, is considerably more involved technically. An obvious analogue of Corollary 1 holds here when $e_i(t) \to 0$ as $t \to \infty$.

Similar to (1.15) we consider the special linear case of (1.16) defined by

$$h_i(t, x, y, z) \equiv h_i > 0, \qquad f(x) = x, \qquad e_i(t) \equiv 0, \qquad (1.20)$$

where the h_i are constants. Lemma 4.1 shows that the real part of each of the characteristic roots of the coefficient matrix associated with (1.16, 1.20) is negative if either (1.17) or (1.18) is satisfied. The remarks following (1.15) are equally applicable here.

On comparing Theorems 2 and 3 it might be suspected that more than one y type component could appear in (1.16) without causing any change in the conclusion of Theorem 3. By means of an example, Remark 4.1, it is shown that this is not the case.

2. Proof of Theorem 1

Lemma 2.1. *Let* (1.2, 1.7, 1.9) *as well as*

$$0 \leqq h(t, x, z) \qquad (0 \leqq t < \infty;\ -\infty < x, z < \infty) \tag{2.1}$$

$$f(x) \in C(-\infty, \infty) \tag{2.2}$$

$$x f(x) \geqq 0 \qquad (-\infty < x < \infty) \tag{2.3}$$

be satisfied. Then given any $x_0, \dot{x}_0$, *there exists a positive constant* $K = K(x_0, \dot{x}_0)$ *such that the solution* $x(t)$ *of* (1.1) *satisfying* $x(0) = x_0$, $\dot{x}(0) = \dot{x}_0$ *exists on* $0 \leqq t < \infty$ *and satisfies*

$$|x(t)|, |\dot{x}(t)| \leqq K \qquad (0 \leqq t < \infty). \tag{2.4}$$

Proof. Equation (1.1) is equivalent to the system

$$\begin{aligned} \dot{x} &= -z, \\ \dot{z} &= -h(t, x, -z)\, z + f(x) - e(t). \end{aligned} \tag{2.5}$$

The present notation is used, rather than the more conventional $\dot{x} = z$, only because it ties in better with (1.11, 1.16) and in turn with the applications of Section 5.

Let x_0, z_0 be given $(z_0 = -\dot{x}_0)$. Let $x = x(t)$, $z = z(t)$ denote the unique solution of (2.5) satisfying $x(0) = x_0$, $z(0) = z_0$ in the remainder of this and the following two lemmas. Then there exists a $t_1 = t_1(x_0, z_0) > 0$ such that $x(t)$, $z(t)$ exists on $0 \leqq t \leqq t_1$.

Define

$$E(t) = g\big(x(t)\big) + \tfrac{1}{2} z^2(t) \qquad (0 \leqq t \leqq t_1), \tag{2.6}$$

where g is defined in (1.7). Differentiating (2.6) and using (2.5) yields

$$\dot{E}(t) = -h\big(t, x(t), -z(t)\big)\, z^2(t) - e(t)\, z(t). \tag{2.7}$$

By (2.1) this implies

$$\dot{E}(t) \leqq -e(t)\, z(t) \leqq |e(t)|\, |z(t)| \leqq |e(t)| + 2|e(t)|\, E(t),$$

which on integrating yields

$$E(t) \leqq E(0) + \int_0^\infty |e(\tau)|\, d\tau + 2\int_0^t |e(\tau)|\, E(\tau)\, d\tau \qquad (0 \leqq t \leqq t_1).$$

Applying the Gronwall inequality, see [*1*, p. 37], one has

$$E(t) \leqq \left\{E(0) + \int_0^\infty |e(\tau)|\, d\tau\right\} \exp\left[2\int_0^\infty |e(\tau)|\, d\tau\right] \qquad (0 \leqq t \leqq t_1). \tag{2.8}$$

Upon noting that the right-hand side of (2.8) is finite and independent of t_1, the conclusion follows readily from (1.7, 2.6) and a routine continuation argument.

Lemma 2.2. *Let the hypothesis of Lemma* 2.1 *as well as* (1.3, 1.4, 1.8) *be satisfied. Then the solution* $x(t)$ *of* (1.1) *of Lemma* 2.1 *satisfies*

$$\lim_{t \to \infty} \dot{x}(t) = 0. \tag{2.9}$$

Proof. In the notation of Lemma 2.1, (2.9) is equivalent to the assertion

$$\lim_{t\to\infty} z(t) = 0. \tag{2.10}$$

Integrating (2.7) and using (1.3) yields

$$0 \leqq E(t) \leqq E(0) - k\int_0^t z^2(\tau)\,d\tau - \int_0^t e(\tau)\,z(\tau)\,d\tau. \tag{2.11}$$

By (1.9, 2.4) there exists a $K = K(x_0, z_0)$ such that

$$E(0) + \left\{\sup_{0\leqq t<\infty} |z(t)|\right\}\int_0^\infty |e(\tau)|\,d\tau = K < \infty, \tag{2.12}$$

and (2.11, 2.12) readily imply

$$\int_0^\infty z^2(\tau)\,d\tau < \infty. \tag{2.13}$$

However, (1.4, 1.8, 2.2, 2.4) and the second equation of (2.5) imply the existence of a $K = K(x_0, z_0)$ such that

$$\left|\frac{d}{dt}\,z^2(t)\right| \leqq K < \infty \qquad (0 \leqq t < \infty). \tag{2.14}$$

By an elementary lemma, see [3, p. 273], (2.10) is an immediate consequence of (2.13, 2.14).

Lemma 2.3. *Let the hypothesis of Theorem 1 be satisfied. Then the solution $x(t)$ of (1.1) of Lemma 2.1 satisfies*

$$\lim_{t\to\infty} x(t) = 0. \tag{2.15}$$

Proof. From (2.4, 2.6) it follows that

$$0 \leqq \overline{\lim_{t\to\infty}}\, E(t) = E_\infty < \infty. \tag{2.16}$$

We now show that the assumption

$$E_\infty > 0 \tag{2.17}$$

leads to a contradiction.

There exists a sequence $\{t_i\}$ such that

$$\lim_{n\to\infty} t_n = \infty, \tag{2.18}$$

$$\lim_{n\to\infty} E(t_n) = E_\infty. \tag{2.19}$$

From (2.6, 2.10, 2.18, 2.19) one easily has

$$\lim_{n\to\infty} g\left(x(t_n)\right) = E_\infty. \tag{2.20}$$

The condition (1.6) implies that $g(x)$ is strictly decreasing on $-\infty < x \leqq 0$ and strictly increasing on $0 \leqq x < \infty$. Hence, by (1.7, 2.16, 2.17, 2.20) there exists a finite

$$\bar{x} \neq 0 \tag{2.21}$$

and a subsequence of $\{t_i\}$, which without loss of generality we may take to be $\{t_i\}$, such that

$$\lim_{n\to\infty} x(t_n) = \bar{x}. \tag{2.22}$$

Integrating the second equation of (2.5) yields

$$\begin{aligned} z(t_n+1) - z(t_n) = &-\int_{t_n}^{t_n+1} h\big(\tau, x(\tau), -z(\tau)\big)\, z(\tau)\, d\tau + f\big(x(t_n)\big) + \\ &+\int_{t_n}^{t_n+1} [f(x(\tau)) - f(x(t_n))]\, d\tau - \int_{t_n}^{t_n+1} e(\tau)\, d\tau. \end{aligned} \tag{2.23}$$

From (1.4, 1.5, 2.4, 2.23) it follows that there exists a finite $K=K(x_0, z_0)$ such that

$$|f(x(t_n))| \le |z(t_n+1) - z(t_n)| + K\int_{t_n}^{t_n+1} |z(\tau)|\, d\tau + K\int_{t_n}^{t_n+1} |x(\tau) - x(t_n)|\, d\tau + \int_{t_n}^{t_n+1} |e(\tau)|\, d\tau,$$

and, hence,

$$|f(x(t_n))| \le (2+K) \sup_{t_n \le t < \infty} |z(t)| + K\int_{t_n}^{t_n+1} |x(\tau) - x(t_n)|\, d\tau + \int_{t_n}^{\infty} |e(\tau)|\, d\tau. \tag{2.24}$$

From the first equation of (2.5) we have

$$x(t) - x(t_n) = -\int_{t_n}^{t} z(\tau)\, d\tau,$$

so that

$$|x(t) - x(t_n)| \le \sup_{t_n \le t < \infty} |z(t)| \qquad (t_n \le t \le t_n + 1). \tag{2.25}$$

From (2.24, 2.25) one has

$$|f(x(t_n))| \le 2(1+K) \sup_{t_n \le t < \infty} |z(t)| + \int_{t_n}^{\infty} |e(\tau)|\, d\tau. \tag{2.26}$$

Letting $n\to\infty$ in (2.26) and using (1.9, 2.10, 2.18) yields

$$\lim_{n\to\infty} f\big(x(t_n)\big) = 0. \tag{2.27}$$

It follows from the continuity of f and (2.22, 2.27) that

$$f(\bar{x}) = 0. \tag{2.28}$$

However, (2.21, 2.28) contradict (1.6) and, thus, (2.17) is impossible. From this it follows that $E_\infty = 0$, which easily yields (2.15).

The preceding lemmas establish Theorem 1. It may be noted that Lemma 2.3 can be proved by a direct argument using essentially the calculations of (2.23) to (2.28); however, the above proof is more like that of Section 4.

Remark 2.1. As already noted in Section 1, the equation $\ddot{x} + x = 0$ shows that (1.3) cannot be dropped entirely. It is not sufficient to replace (1.3) by $h > 0$, as the following shows.

Let $h(t) \ge 0$ and $h(t) \in L_1(0, \infty)$. Then there exist solutions of

$$\ddot{x} + h(t)\,\dot{x} + x = 0 \tag{2.29}$$

which do not tend to zero as $t\to\infty$.

Proof. The change of variables $x=\xi+\sin t$ transforms (2.29) into

$$\ddot{\xi}+h(t)\,\dot{\xi}+\xi=-h(t)\cos t,$$

which is equivalent to the system

$$\dot{\xi}=-\eta,\qquad \dot{\eta}=-h(t)\,\eta+\xi+h(t)\cos t. \tag{2.30}$$

By hypothesis there exists a finite $t_1>0$ such that

$$\left\{\int_{t_1}^{\infty} h(\tau)\,d\tau\right\}\exp\left[2\int_{t_1}^{\infty} h(\tau)\,d\tau\right]\leqq \tfrac{1}{8}. \tag{2.31}$$

Consider the solution $\xi_1(t)$, $\eta_1(t)$ of (2.30) that is defined by $\xi_1(t_1)=\eta_1(t_1)=0$. Define

$$E(t)=\tfrac{1}{2}\xi_1^2(t)+\tfrac{1}{2}\eta_1^2(t).$$

Following Lemma 2.1 and using (2.31), one obtains

$$\dot{E}(t)=-h(t)\,\eta_1^2(t)+\eta_1(t)\,h(t)\cos t\leqq h(t)+2h(t)\,E(t),$$

$$E(t)\leqq\int_{t_1}^{\infty} h(\tau)\,d\tau+2\int_{t_1}^{t} h(\tau)\,E(\tau)\,d\tau\qquad (t_1\leqq t<\infty),$$

$$E(t)\leqq\tfrac{1}{8},\qquad |\xi_1(t)|\leqq\tfrac{1}{2}\qquad (t_1\leqq t<\infty).$$

Hence, the solution $x_1(t)=\xi_1(t)+\sin t$ of (2.29) does not tend to zero as $t\to\infty$, which completes the proof.

Remark 2.2. The following shows that (1.4) *cannot be dropped entirely. Let* $\dot{h}(t)\in C(0,\infty)$, $h(t)>0$, $h^{-1}(t)\in L_1(0,\infty)$, *and* $[\dot{h}(t)+1]\,h^{-2}(t)\in L_1(0,\infty)$. *Then there exist solutions of* (2.29) *which do not tend to zero as* $t\to\infty$.

Proof. The change of variables $x=\xi+\varphi(t)$, where

$$\varphi(t)=\exp\left[-\int_0^t\frac{1}{h(\tau)}\,d\tau\right], \tag{2.32}$$

transforms (2.29) into

$$\ddot{\xi}+h(t)\,\dot{\xi}+\xi=-\ddot{\varphi}(t),$$

which is equivalent to the system

$$\dot{\xi}=-\eta,\qquad \dot{\eta}=-h(t)\,\eta+\xi+\ddot{\varphi}(t). \tag{2.33}$$

From (2.32) and the hypothesis there exists a constant λ such that

$$0<\lambda<\varphi(t)\qquad (0\leqq t<\infty). \tag{2.34}$$

As

$$\ddot{\varphi}(t)=\frac{1}{h^2(t)}\,[\dot{h}(t)+1]\exp\left[-\int_0^t\frac{1}{h(\tau)}\,d\tau\right],$$

it follows from the hypothesis that there exists a finite $t_1>0$ such that

$$\left\{\int_{t_1}^{\infty}|\ddot{\varphi}(t)|\,dt\right\}\exp\left[2\int_{t_1}^{\infty}|\ddot{\varphi}(t)|\,dt\right]\leqq\frac{\lambda^2}{8}.$$

Consider the solution $\xi_1(t)$, $\eta_1(t)$ of (2.33) that is defined by $\xi_1(t_1)=\eta_1(t_1)=0$. Exactly as in Remark 2.1 we now obtain

$$|\xi_1(t)| \leqq \frac{\lambda}{2} \qquad (t_1 \leqq t < \infty). \tag{2.35}$$

Hence, the solution $x_1(t)=\xi_1(t)+\varphi(t)$ of (2.29) does not tend to zero as $t\to\infty$, as is evident from (2.34, 2.35).

3. Proof of Theorem 2

First we consider the case that (1.12) is satisfied. Let x_0, z_0 be given. Let $x(t)$, $z(t)$ denote the unique solution of (1.11) satisfying $x(0)=x_0$, $z(0)=z_0$ Then $x(t)$, $z(t)$ exists for t sufficiently small Define

$$E(t)=\frac{1}{c}g\left(x(t)\right)+\frac{1}{2}\sum_{i=1}^{n} z_i^2(t). \tag{3.1}$$

As in Lemma 2.1 we now obtain

$$\begin{aligned} \dot{E}(t) &= -\sum_{i=1}^{n} h_i\left(t, x(t), z(t)\right) z_i^2(t) + \sum_{i=1}^{n} e_i(t)\, z_i(t), \\ \dot{E}(t) &\leq \sum_{i=1}^{n} |e_i(t)| + 2\sum_{i=1}^{n} |e_i(t)|\, E(t), \end{aligned} \tag{3.2}$$

and

$$E(t) \leqq \left\{E(0)+\int_0^\infty \sum_{i=1}^{n} |e_i(\tau)|\, d\tau\right\} \exp\left[2\int_0^\infty \sum_{i=1}^{n} |e_i(\tau)|\, d\tau\right],$$

which yields the existence and boundedness of $x(t)$, $z(t)$ on $0\leqq t<\infty$.

Following Lemma 2.2, one obtains

$$0 \leqq E(t) \leqq E(0) - k\int_0^t \sum_{i=1}^{n} z_i^2(\tau)\, d\tau + \int_0^t \sum_{i=1}^{n} e_i(\tau)\, z_i(\tau)\, d\tau \tag{3.3}$$

from (1.3, 3.2), and

$$E(0)+\left\{\sup_{0<t<\infty} \sum_{i=1}^{n} |z_i(t)|\right\} \int_0^\infty \sum_{i=1}^{n} |e_i(\tau)|\, d\tau = K < \infty \tag{3.4}$$

from (1.9) and the boundedness of $x(t)$, $z(t)$, where $K=K(x_0, z_0)$. However, (3.3, 3.4) imply

$$\int_0^\infty \sum_{i=1}^{n} z_i^2(\tau)\, d\tau < \infty. \tag{3.5}$$

As before there exists a $K=K(x_0, z_0)$ such that

$$\left|\frac{d}{dt}\sum_{i=1}^{n} z_i^2(t)\right| \leq K < \infty \qquad (0 \leq t < \infty). \tag{3.6}$$

From (3.5, 3.6) it follows that

$$\lim_{t\to\infty} z(t) = 0,$$

which is the second half of (1.14).

The proof of Lemma 2.3 is easily adapted to obtain the first half of (1.14). One simply applies the argument of Lemma 2.3 to the first and $j+1^{\text{st}}$ equations of (1.11) (here is where $b_j \neq 0$ is needed). This completes the proof of Theorem 2 under hypothesis (1.12).

Under hypothesis (1.13) the proof goes much the same way. One now defines

$$E(t)=g\big(x(t)\big)+\frac{1}{2}\sum_{i=1}^{n}\frac{a_i}{b_i}z_i^2(t), \tag{3.7}$$

and obtains analogous formulae, for example

$$E(t)\leqq\Big\{E(0)+\int_0^\infty\sum_{i=1}^{n}\frac{a_i}{b_i}\,|e_i(\tau)|\,d\tau\Big\}\exp\Big[2\int_0^\infty\sum_{i=1}^{n}|e_i(\tau)|\,d\tau\Big],$$

$$\int_0^\infty\sum_{i=1}^{n}\frac{a_i}{b_i}z_i^2(\tau)\,d\tau<\infty,\qquad\Big|\frac{d}{dt}\sum_{i=1}^{n}\frac{a_i}{b_i}z_i^2(t)\Big|\leqq K<\infty\qquad(0\leqq t<\infty).$$

The adaptation of the proof of Lemma 2.3 is exactly the same as the one mentioned above, except that here any one of the last n equations of (1.11) may be used. This completes the proof of Theorem 2.

The special case of (1.11) defined by (1.15) may be written as

$$\frac{d}{dt}\begin{pmatrix}x\\ z_1\\ z_2\\ \vdots\\ z_n\end{pmatrix}=\begin{pmatrix}0&-a_1&-a_2&\dots&-a_n\\ b_1&-h_1&0&\dots&0\\ b_2&0&-h_2&\dots&0\\ \cdot&\cdot&\cdot&\cdot&\cdot\\ b_n&0&0&\dots&-h_n\end{pmatrix}\begin{pmatrix}x\\ z_1\\ z_2\\ \vdots\\ z_n\end{pmatrix}, \tag{3.8}$$

where the $h_i>0$. We denote the coefficient matrix in (3.8) by G.

Lemma 3.1. *If either* (1.12) *or* (1.13) *is satisfied, then all the roots of* $\det(\lambda I-G)=0$ *lie in the left half plane* $\operatorname{Re}\lambda<0$.

Proof. It may be established by induction that

$$p(\lambda)=\frac{\det(\lambda I-G)}{\prod_{k=1}^{n}(\lambda+h_k)}=\lambda+\sum_{k=1}^{n}\frac{a_k b_k}{\lambda+h_k}.$$

As the $h_k>0$, it is clearly sufficient to show that $p(\lambda)\neq 0$ for $\operatorname{Re}\lambda\geqq 0$. Let $\lambda=\sigma+i\tau$. Then

$$\operatorname{Re}p(\lambda)=\sigma+\sum_{k=1}^{n}\frac{a_k b_k(\sigma+h_k)}{(\sigma+h_k)^2+\tau^2}. \tag{3.9}$$

From (3.9) it is evident that $\operatorname{Re}p(\lambda)>0$ for $\sigma\geqq 0$ if either (1.12) or (1.13) is satisfied, which completes the proof.

4. Proof of Theorem 3

First we consider the case that (1.17) is satisfied. Let x_0, y_0, z_0 be given. Let $x(t)$, $y(t)$, $z(t)$ denote the unique solution of (1.16) satisfying $x(0)=x_0$, $y(0)=y_0$, $z(0)=z_0$. Then $x(t)$, $y(t)$, $z(t)$ exists for t sufficiently small. By the same methods employed in the proof of Theorem 2, one can now obtain the existence of $x(t)$, $y(t)$, $z(t)$ on $0\leqq t<\infty$ and

$$|x(t)|,\ |y(t)|,\ |z(t)|\leqq K<\infty\qquad(0\leqq t<\infty), \tag{4.1}$$

$$\lim_{t\to\infty}z(t)=0, \tag{4.2}$$

where $K=K(x_0, y_0, z_0)$. The analogous formulae are

$$E(t) = \frac{1}{c}\, g\left(x(t)\right) + \frac{1}{2}\, y^2(t) + \frac{1}{2} \sum_{i=1}^{n} z_i^2(t) \tag{4.3}$$

and, for example

$$E(t) \leqq \left\{E(0) + \int_0^{\infty} \sum_{i=0}^{n} |e_i(\tau)|\, d\tau\right\} \exp\left[2 \int_0^{\infty} \sum_{i=0}^{n} |e_i(\tau)|\, d\tau\right],$$

$$0 \leqq E(t) \leqq E(0) - k \int_0^t \sum_{i=1}^{n} z_i^2(\tau)\, d\tau + \int_0^t e_0(\tau)\, y(\tau)\, d\tau + \int_0^t \sum_{i=1}^{n} e_i(\tau)\, z_i(\tau)\, d\tau,$$

$$\int_0^{\infty} \sum_{i=1}^{n} z_i^2(\tau)\, d\tau < \infty, \qquad \left|\frac{d}{dt} \sum_{i=1}^{n} z_i^2(t)\right| \leqq K < \infty \qquad (0 \leqq t < \infty).$$

Thus, it remains to show that

$$\lim_{t\to\infty} x(t) = 0, \qquad \lim_{t\to\infty} y(t) = 0. \tag{4.4}$$

From (4.1, 4.3) it follows that

$$0 \leqq \overline{\lim_{t\to\infty}}\, E(t) = E_\infty < \infty. \tag{4.5}$$

We now show that the assumption

$$E_\infty > 0 \tag{4.6}$$

leads to a contradiction.

There exists a sequence $\{t_i\}$ such that

$$\lim_{n\to\infty} t_n = 0 \tag{4.7}$$

$$\lim_{n\to\infty} E(t_n) = E_\infty. \tag{4.8}$$

From (4.2, 4.3, 4.7, 4.8) one easily has

$$\lim_{n\to\infty} \left\{\frac{1}{c}\, g\left(x(t_n)\right) + \frac{1}{2}\, y^2(t_n)\right\} = E_\infty. \tag{4.9}$$

Consider the system

$$\dot{x} = -c\, b_0\, y, \qquad \dot{y} = b_0\, f(x), \tag{4.10}$$

which is equivalent to the equation

$$\ddot{x} + c\, b_0^2\, f(x) = 0. \tag{4.11}$$

It is well known that under hypotheses (1.5, 1.6, 1.7) (and $c\, b_0^2 > 0$) all solutions of (4.11), and hence of (4.10), are periodic. Let $x=\varphi(t)$, $y=\psi(t)$ be the solution of (4.10) such that $\varphi(0)=0$, $\psi(0)=\sqrt{2E_\infty}$. Let $\Omega>0$ be the smallest positive period of $\varphi(t)$, $\psi(t)$. As (4.10) is autonomous, $x=\varphi(t+\varrho)$, $y=\psi(t+\varrho)$ is a solution of (4.10) for all real ϱ. Furthermore, these and only these solutions of (4.10) have the same orbit, $\mathscr{C}$, in the x, y plane. The graph of $\mathscr{C}$ is given by

$$\frac{1}{c}\, g(x) + \frac{1}{2}\, y^2 = E_\infty. \tag{4.12}$$

From the nonconstancy, a consequence of (4.6), and periodicity of $\psi(t)$ it is easily seen that there exists a positive constant μ such that

(4.13) $$\min_{-\infty<t_1<\infty} \max_{t_1 \leqq t_2 \leqq t_1+\Omega} |\psi(t_2) - \psi(t_1)| = \mu > 0.$$

Let $\varepsilon > 0$ be given. It is readily seen from (1.9, 4.2, 4.7, 4.9, 4.12), the geometry of the curves

$$\frac{1}{c} g(x) + \frac{1}{2} y^2 = \text{const.} \geqq 0,$$

and the preceding discussion that there exists a positive integer N_ε and a real number $\varrho_\varepsilon \geqq 0$ such that

(4.14) $$|x(t_{N_\varepsilon}) - \varphi(t_{N_\varepsilon} + \varrho_\varepsilon)| \leqq \varepsilon, \quad |y(t_{N_\varepsilon}) - \psi(t_{N_\varepsilon} + \varrho_\varepsilon)| \leqq \varepsilon,$$

(4.15) $$|z(t)| \leqq \varepsilon \qquad (t_{N_\varepsilon} \leqq t < \infty),$$

(4.16) $$\int_{t_{N_\varepsilon}}^{\infty} |e_i(t)|\, dt \leqq \varepsilon \qquad (i = 0, \ldots, n).$$

Define

(4.17) $$\tilde{\varphi}(t) = \tilde{\varphi}(t, \varepsilon) = \varphi(t + \varrho_\varepsilon), \quad \tilde{\psi}(t) = \tilde{\psi}(t, \varepsilon) = \psi(t + \varrho_\varepsilon).$$

As (4.10) is autonomous, one has from (1.16, 4.10)

$$\frac{d}{dt}[x(t) - \tilde{\varphi}(t)] = -c\, b_0 [y(t) - \tilde{\psi}(t)] - c \sum_{i=1}^{n} b_i z_i(t)$$
$$\frac{d}{dt}[y(t) - \tilde{\psi}(t)] = b_0 [f(x(t)) - f(\tilde{\varphi}(t))] + e_0(t),$$

which upon integrating yields

(4.18) $$\begin{aligned} x(t) - \tilde{\varphi}(t) &= x(t_{N_\varepsilon}) - \tilde{\varphi}(t_{N_\varepsilon}) - c\, b_0 \int_{t_{N_\varepsilon}}^{t} [y(\tau) - \tilde{\psi}(\tau)]\, d\tau - c \sum_{i=1}^{n} b_i \int_{t_{N_\varepsilon}}^{t} z_i(\tau)\, d\tau, \\ y(t) - \tilde{\psi}(t) &= y(t_{N_\varepsilon}) - \tilde{\psi}(t_{N_\varepsilon}) + b_0 \int_{t_{N_\varepsilon}}^{t} [f(x(\tau)) - f(\tilde{\varphi}(\tau))]\, d\tau + \int_{t_{N_\varepsilon}}^{t} e_0(\tau)\, d\tau. \end{aligned}$$

From (1.5, 4.1, 4.14, 4.15, 4.16, 4.18) it follows that

(4.19) $$\left.\begin{aligned} |x(t) - \tilde{\varphi}(t)| &\leqq \Big\{1 + c\,\Omega \sum_{i=1}^{n} |b_i|\Big\} \varepsilon + c|b_0| \int_{t_{N_\varepsilon}}^{t} |y(\tau) - \tilde{\psi}(\tau)|\, d\tau \\ |y(t) - \tilde{\psi}(t)| &\leqq 2\varepsilon + K_1 |b_0| \int_{t_{N_\varepsilon}}^{t} |x(\tau) - \tilde{\varphi}(\tau)|\, d\tau, \end{aligned}\right\} \quad (t_{N_\varepsilon} \leqq t \leqq t_{N_\varepsilon} + \Omega)$$

where $K_1 = K_1(x_0, y_0, z_0) > 0$ and, in particular, is independent of ε and n. On adding the two inequalities of (4.19), there results

$$|x(t) - \tilde{\varphi}(t)| + |y(t) - \tilde{\psi}(t)| \leqq K_2 \varepsilon + K_2 \int_{t_{N_\varepsilon}}^{t} \{|x(\tau) - \tilde{\varphi}(\tau)| + |y(\tau) - \tilde{\psi}(\tau)|\}\, d\tau$$
$$(t_{N_\varepsilon} \leqq t \leqq t_{N_\varepsilon} + \Omega),$$

where $K_2 = K_2(x_0, y_0, z_0)$, which by the Gronwall inequality implies

$$|x(t) - \tilde{\varphi}(t)| + |y(t) - \tilde{\psi}(t)| \leqq K_2 \exp[K_2 \Omega]\, \varepsilon \qquad (t_{N_\varepsilon} \leqq t \leqq t_{N_\varepsilon} + \Omega).$$

Thus, for some $K_3 = K_3(x_0, y_0, z_0)$,

$$|y(t) - \tilde{\psi}(t)| \leqq K_3 \varepsilon \qquad (t_{N_\varepsilon} \leqq t \leqq t_{N_\varepsilon} + \Omega). \tag{4.20}$$

The second and $j+2^{\text{nd}}$ equations of (1.16) together with (1.17) imply

$$\dot{y}(t) = \frac{b_0}{b_j} \dot{z}_j(t) + \frac{b_0}{b_j} h_j(t, x(t), y(t), z(t))\, z_j(t) - \frac{b_0}{b_j} e_j(t) + e_0(t),$$

and, hence,

$$\begin{aligned} y(t) - y(t_{N_\varepsilon}) &= \frac{b_0}{b_j} [z_j(t) - z_j(t_{N_\varepsilon})] + \\ &+ \frac{b_0}{b_j} \int_{t_{N_\varepsilon}}^{t} h_j(\tau, x(\tau), y(\tau), z(\tau)) \cdot z_j(\tau)\, d\tau - \frac{b_0}{b_j} \int_{t_{N_\varepsilon}}^{t} e_j(\tau)\, d\tau + \int_{t_{N_\varepsilon}}^{t} e_0(\tau)\, d\tau. \end{aligned} \tag{4.21}$$

From (1.4, 4.1, 4.15, 4.16, 4.21) one obtains

$$|y(t) - y(t_{N_\varepsilon})| \leqq K_4 \varepsilon \qquad (t_{N_\varepsilon} \leqq t \leqq t_{N_\varepsilon} + \Omega), \tag{4.22}$$

for some $K_4 = K_4(x_0, y_0, z_0)$. However, from

$$\tilde{\psi}(t) - \tilde{\psi}(t_{N_\varepsilon}) = \tilde{\psi}(t) - y(t) + y(t) - y(t_{N_\varepsilon}) + y(t_{N_\varepsilon}) - \tilde{\psi}(t_{N_\varepsilon})$$

and (4.17, 4.20, 4.22) it follows that

$$\max_{t_{N_\varepsilon} + \varrho_\varepsilon \leqq t \leqq t_{N_\varepsilon} + \varrho_\varepsilon + \Omega} |\psi(t) - \psi(t_{N_\varepsilon} + \varrho_\varepsilon)| \leqq (2K_3 + K_4)\, \varepsilon. \tag{4.23}$$

Comparing (4.13, 4.23) yields $\mu \leqq (2K_3 + K_4)\,\varepsilon$, which, as ε may be chosen arbitrarily small, contradicts $\mu > 0$.

Thus, it follows that $E_\infty = 0$, which easily implies (4.4) and completes the proof under hypothesis (1.17).

The proof under hypothesis (1.18) is essentially the same as under (1.17). Instead of (4.3) one defines

$$E(t) = g(x(t)) + \frac{1}{2} \frac{a_0}{b_0} y^2(t) + \frac{1}{2} \sum_{i=1}^{n} \frac{a_i}{b_i} z_i^2(t),$$

and modifies the calculations along the same lines as in the formulae immediately following (3.7). This completes the proof of Theorem 3.

The special case of (1.16) defined by (1.20) may be written as

$$\frac{d}{dt} \begin{pmatrix} x \\ y \\ z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix} = \begin{pmatrix} 0 & -a_0 & -a_1 & -a_2 & \dots & -a_n \\ b_0 & 0 & 0 & 0 & \dots & 0 \\ b_1 & 0 & -h_1 & 0 & \dots & 0 \\ b_2 & 0 & 0 & -h_2 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ b_n & 0 & 0 & 0 & \dots & -h_n \end{pmatrix} \begin{pmatrix} x \\ y \\ z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix} \tag{4.24}$$

where the $h_i > 0$. We also denote the coefficient matrix in (4.24) by G.

Lemma 4.1. *If either* (1.17) *or* (1.18) *is satisfied, then all the roots of* $\det(\lambda I - G) = 0$ *lie in the left half plane* $\operatorname{Re}\lambda < 0$.

Proof. It may be established by induction that

$$p(\lambda) = \frac{\det(\lambda I - G)}{\prod_{k=1}^{n}(\lambda + h_k)} = \lambda^2 + a_0 b_0 + \sum_{k=1}^{n} a_k b_k \frac{\lambda}{\lambda + h_k}.$$

As the $h_k > 0$ it is clearly sufficient to show that $p(\lambda) \neq 0$ for $\operatorname{Re}\lambda \geqq 0$. Let $\lambda = \sigma + i\tau$. Then

$$\begin{aligned} \operatorname{Re} p(\lambda) &= \sigma^2 - \tau^2 + a_0 b_0 + \sum_{k=1}^{n} a_k b_k \frac{\sigma(\sigma + h_k) + \tau^2}{(\sigma + h_k)^2 + \tau^2} \\ \operatorname{Im} p(\lambda) &= 2\sigma\tau + \tau \sum_{k=1}^{n} a_k b_k \frac{h_k}{(\sigma + h_k)^2 + \tau^2}. \end{aligned} \tag{4.25}$$

From (4.25) it is easily seen that $\operatorname{Re} p(\lambda)$ and $\operatorname{Im} p(\lambda)$ cannot vanish simultaneously for $\sigma \geqq 0$ if either (1.17) or (1.18) is satisfied, which completes the proof.

Remark 4.1. That more than one y type component cannot appear in (1.16) without changing the conclusion of Theorem 3 is shown by:

Let $f(x)$ *satisfy* (1.5, 1.6, 1.7). *Then there exist solutions of*

$$\dot{x} = -y_1 - y_2 - z, \quad \dot{y}_1 = f(x), \quad \dot{y}_2 = f(x), \quad \dot{z} = -z + f(x) \tag{4.26}$$

which do not tend to zero as $t \to \infty$.

Proof. Let $\tilde{x}(t)$, $\tilde{\xi}(t)$, $\tilde{z}(t)$ be a solution of

$$\dot{x} = -\xi - z, \quad \dot{\xi} = 2f(x), \quad \dot{z} = -z + f(x).$$

Then by Theorem 3 (hypothesis (1.18))

$$\lim_{t \to \infty} \left(\tilde{x}(t), \tilde{\xi}(t), \tilde{z}(t)\right) = (0, 0, 0).$$

However,

$$x^*(t) = \tilde{x}(t), \quad y_1^*(t) = \tfrac{1}{2}\tilde{\xi}(t) - 1, \quad y_2^*(t) = \tfrac{1}{2}\tilde{\xi}(t) + 1, \quad z^*(t) = z(t)$$

s easily shown to be a solution of (4.26), and as

$$\lim_{t \to \infty} \left(x^*(t), y_1^*(t), y_2^*(t), z^*(t)\right) = (0, -1, 1, 0)$$

the proof is complete.

A little reflection shows that the key to the preceding example is the existence of the critical points $x = 0$, $y_1 = \pm\sqrt{k}$, $y_2 = \mp\sqrt{k}$ on

$$g(x) + \tfrac{1}{2} y_1^2 + \tfrac{1}{2} y_2^2 = k > 0,$$

which is an integral for the system

$$\dot{x} = -y_1 - y_2, \quad \dot{y}_1 = f(x), \quad \dot{y}_2 = f(x).$$

5. Applications to Reactor Dynamics

In this section Theorem 2 with $n=1$ (which is essentially Theorem 1) and Theorem 3 are applied to certain problems in reactor dynamics. For earlier work in this direction as well as some references to the physical literature see [5].

The dynamic equations for a class of homogeneous nuclear reactors are of the form

$$\dot{u} = -\alpha T, \quad \varepsilon \dot{T} = -h(t, u, T)\, T + f(u) + e(t). \tag{5.1}$$

In the case of NEWTON's law of cooling studied previously [5] one has

$$h(t, u, T) \equiv 1, \quad f(u) = \exp[u] - 1, \quad e(t) \equiv 0, \tag{5.2}$$

with the physical interpretations

$$\begin{aligned} u &= \text{logarithm of the ratio of the reactor power to the stationary power } (u \equiv 0 \text{ at equilibrium}),\\ T &= \text{deviation of the temperature from the equilibrium temperature } (T \equiv 0 \text{ at equilibrium}),\\ \varepsilon &= \text{heat capacity},\\ -\alpha &= \text{ratio of the temperature coefficient of reactivity to the mean life of neutrons.} \end{aligned} \tag{5.3}$$

In (5.1) h is a heat conduction term and e is a forcing term.

An immediate consequence of Theorem 2 ($n=1$) is

Theorem 4. *Let the functions h, f, e satisfy the hypothesis of Theorem 1, and let the constants α, ε satisfy $\varepsilon > 0$, $\alpha > 0$. Then given any u_0, T_0, the solution $u(t)$, $T(t)$ of (5.1) satisfying $u(0) = u_0$, $T(0) = T_0$ exists on $0 \leqq t < \infty$, and*

$$\lim_{t\to\infty} u(t) = 0, \quad \lim_{t\to\infty} T(t) = 0. \tag{5.4}$$

It may be noted that Theorem 4 restricted to (5.1, 5.2) generalizes Theorem 3.1 of [5] in that under the same hypothesis it establishes global rather than local asymptotic stability. Further, while the earlier theorem asserted local asymptotic stability, the proof was quite incomplete and all that was really established was ordinary local stability. Comparing Theorem 4 itself with (5.2), one notes that the only relevant property of $f = \exp[u] - 1$ is that it represents a "spring" in the usual sense of (1.5, 1.6, 1.7), and that of $h = 1$ is that it represents a "positive damping" in the sense of (1.3, 1.4).

The dynamic equations for a class of heterogeneous reactors of m ($m \geqq 2$) media are of the form

$$\begin{aligned} \dot{u} &= -\sum_{j=1}^{m} \alpha_j T_j,\\ \varepsilon_i \dot{T}_i &= -h(t, u, T) \sum_{j=1}^{m} X_{ij}(T_i - T_j) + \eta_i f(u) + e_i(t) \qquad (i = 1, \ldots, m), \end{aligned} \tag{5.5}$$

where $T = \mathrm{col}(T_1, \ldots, T_m)$. The special case studied in [5] had

$$h(t, u, T) \equiv 1, \quad f(u) = \exp[u] - 1, \quad e_i(t) \equiv 0 \qquad (i = 1, \ldots, m), \tag{5.6}$$

with the physical interpretations

$$
\begin{aligned}
u &= \text{logarithm of the ratio of the reactor power to the stationary power } (u \equiv 0 \text{ at equilibrium}),\\
T_i &= \text{deviation of the temperature from the equilibrium temperature in the } i^{\text{th}} \text{ medium } (T_i \equiv 0 \text{ at equilibrium}),\\
\varepsilon_i &= \text{heat capacity in the } i^{\text{th}} \text{ medium},\\
-\alpha_i &= \text{ratio of the temperature coefficient of reactivity in the } i^{\text{th}} \text{ medium to the mean life of neutrons},\\
\eta_i &= \text{fraction of the power generated in the } i^{\text{th}} \text{ medium (physically } \textstyle\sum \eta_i = 1),\\
X_{ij} &= \text{thermal conductivity from medium } i \text{ to medium } j \text{ (physically } X_{ij} = X_{ji}).
\end{aligned}
\tag{5.7}
$$

In (5.5) h is a heat conduction term and the e_i are forcing terms.

In order to apply Theorem 3 to (5.5) (it will shortly be seen that Theorem 2 is not applicable here), it is first necessary to transform it to the form (1.16). The notation

$$
\alpha = \operatorname{col}(\alpha_1, \ldots, \alpha_m), \qquad \eta = \operatorname{col}(\eta_1, \ldots, \eta_m), \qquad e = \operatorname{col}(e_1, \ldots, e_m),
$$

$$
\varepsilon = \operatorname{diag}(\varepsilon_1, \ldots, \varepsilon_m), \qquad \mu' = \text{transpose of } \mu \text{ (for any } \mu),
$$

will be used. From (5.5) it is clear that

$$
\varepsilon_i \dot{T}_i = -h(t, u, T)\Big\{\sum_{j \neq i} X_{ij}\Big\} T_i + h(t, u, T) \sum_{j \neq i} X_{ij} T_j + \eta_i f(u) + e_i(t) \qquad (i = 1, \ldots, m). \tag{5.8}
$$

Define the matrix $A = (A_{ij})$ by

$$
A_{ii} = \sum_{j \neq i} X_{ij}, \qquad A_{ij} = -X_{ij} \qquad (i \neq j). \tag{5.9}
$$

Then from (5.8, 5.9) one sees that (5.5) may be written as

$$
\begin{aligned}
\dot{u} &= -T'\alpha,\\
\varepsilon \dot{T} &= -h(t, u, T) A T + \eta f(u) + e(t).
\end{aligned}
\tag{5.10}
$$

We need the following algebraic

Lemma 5.1. *Let*

$$
X_{ij} = X_{ji} > 0 \quad (i \neq j), \qquad \varepsilon_i > 0 \quad (i = 1, \ldots, m). \tag{5.11}
$$

Then (i) *the real matrix A defined by* (5.9) *is symmetric, and precisely one of its characteristic values is zero, while the remaining $m-1$ characteristic values are positive*; (ii) *there exists a real nonsingular matrix R such that*

$$
R' \varepsilon R = I, \qquad R' A R = D = \operatorname{diag}(d_1, \ldots, d_m), \tag{5.12}
$$

where I is the unit m by m matrix. Furthermore, precisely one of the d_i is zero, say $d_j = 0$, while the rest are positive.

Proof. (i) That A is symmetric is obvious from the definition (5.9) and the first half of (5.11). Consider the quadratic form

$$W(\mu) = \mu' A \mu. \tag{5.13}$$

A direct calculation yields

$$W(\mu) = \sum_{i=1}^{m-1} \sum_{j=i+1}^{m} X_{ij}(\mu_i - \mu_j)^2,$$

from which, using the first half of (5.11), it is evident that $W(\mu)$ is positive for all μ except multiples of $\mathrm{col}(1, \ldots, 1)$, for which it vanishes. The statement about the characteristic values of A now follows from well-known facts about real symmetric matrices.

(ii) By a well-known theorem, see [*6*, p. 187], (5.12) is an immediate consequence of the positive definite real symmetric nature of ε and the symmetry of the real matrix A. Making the change of variables $\mu = R\xi$ in (5.13) and using the second half of (5.12), one has

$$W(\mu) = W(R\xi) = \xi' R' A R \xi = \sum_{i=1}^{m} d_i \xi_i^2,$$

which together with (i) easily yields the statement about the d_i.

Letting $T = RQ$ in (5.10) yields

$$\begin{aligned} \dot{u} &= -Q' a \\ \dot{Q} &= -h(t, u, RQ) D Q + b f(u) + \tilde{e}(t) \end{aligned} \tag{5.14}$$

where

$$a = R'\alpha, \qquad b = R'\eta, \qquad \tilde{e}(t) = R' e(t), \tag{5.15}$$

which written in component form is

$$\begin{aligned} \dot{u} &= -a_j Q_j - \sum_{k \neq j} a_k Q_k \\ \dot{Q}_j &= b_j f(u) + \tilde{e}_j(t) \\ \dot{Q}_i &= -h(t, u, RQ) d_i Q_i + b_i f(u) + \tilde{e}_i(t) \qquad (i \neq j). \end{aligned} \tag{5.16}$$

Under hypothesis (5.11) the system (5.14), or equivalently (5.16), has the same stability properties as (5.10) since R is a nonsingular constant matrix. It is immediately seen that (5.16) is of the same form as (1.16). Therefore, we have as a consequence of Theorem 3.

Theorem 5. *Let the functions h, f, e_i be sufficiently smooth for a local existence and uniqueness theorem to hold for* (5.5)*; the h, f, e_i satisfy the hypothesis of Theorem* 1*; the X_{ij} and ε_i satisfy* (5.11)*; R be as defined in Lemma* 5.1*; and a, b be as defined by* (5.15)*. Furthermore, let a, b satisfy either*

(5.17) *$a = cb$, where $c > 0$, and b is arbitrary except that $b_j \neq 0$ and at least one other b_i is not zero,*

or

$$a_i / b_i > 0 \qquad (i = 0, 1, \ldots, m). \tag{5.18}$$

Then given any u_0, T_0, *the solution* $u(t)$, $T(t)$ *of* (5.5) *satisfying* $u(0)=u_0$, $T(0)=T_0$ *exists on* $0 \leqq t < \infty$, *and*

$$\lim_{t\to\infty} u(t) = 0, \qquad \lim_{t\to\infty} T(t) = 0. \tag{5.19}$$

The remarks following Theorem 4 apply equally well here to the comparison of Theorem 5 with Theorem 4.2 of [5]. The verification of hypothesis (5.17) or (5.18) is a purely algebraic matter and does not involve solving any differential equations. R is determined by the given matrices A and ε; and then a, b are defined by (5.15) in terms of the given vectors α, η. Hypothesis (5.17) is analogous to the crucial hypothesis, (2.6) of [7], used in the analysis of the integrodifferential equations for a certain continuous medium reactor.

We note that if

$$\alpha = c\,\eta \qquad (c > 0), \tag{5.20}$$

then from (5.15) it follows immediately that

$$a = c\,b \qquad (c > 0). \tag{5.21}$$

Under condition (5.21) one obtains (this is the first third of the proof of Theorem 3) not (5.19) but merely that $u(t)$, $T(t)$ exist on $0 \leqq t < \infty$ and are bounded there. However, the set of vectors in b space excluded by condition (5.17) is easily seen to be of Lebesgue measure zero. Hence, using the nonsingularity of R, one can give the following modified form of Theorem 5: *If* (5.20) *is satisfied, then for almost all* η (5.19) *holds, and on the exceptional* η *set* $u(t)$, $T(t)$ *are bounded.*

References

[1] Coddington, E. A., & N. Levinson: Theory of Differential Equations. New York (1955).
[2] Antosiewicz, H. A.: On nonlinear differential equations of second order with integrable forcing term. J. London Math. Soc., vol. **30**, 64–67 (1955).
[3] Lefschetz, S.: Differential Equations: Geometric Theory. New York (1957).
[4] Bellman, R.: Stability Theory of Differential Equations. New York 1953.
[5] Ergen, W. K., H. J. Lipkin & J. A. Nohel: Applications of Liapounov's second method in reactor dynamics. Jour. Math. and Physics, vol. **36**, 36–48 (1957).
[6] Perlis, S : Theory of Matrices. Cambridge 1952.
[7] Levin, J. J., & J. A. Nohel: On a system of integrodifferential equations occurring in reactor dynamics. To appear in Jour. Math. and Mech. (1960).

Lincoln Laboratory
Massachusetts Institute of Technology
Cambridge, Massachusetts
and
Georgia Institute of Technology
Atlanta, Georgia

(Received February 3, 1960)

Singular Perturbations of Two-Point Boundary Problems for Systems of Ordinary Differential Equations

W. A. HARRIS, JR.

Communicated by C. C. LIN

Contents

Abstract

Asymptotic solutions of linear systems of ordinary differential equations are employed to discuss the relationship of the solution of a certain "complete" boundary problem.

$$\left\{\begin{aligned} \frac{dx_1}{dt} &= A_{11}(t,\varepsilon)\,x_1(t,\varepsilon) + \cdots + A_{1p}(t,\varepsilon)\,x_p(t,\varepsilon)\\ \varepsilon^{h_2}\frac{dx_2}{dt} &= A_{21}(t,\varepsilon)\,x_1(t,\varepsilon) + \cdots + A_{2p}(t,\varepsilon)\,x_p(t,\varepsilon)\\ &\;\;\vdots\\ \varepsilon^{h_p}\frac{dx_p}{dt} &= A_{p1}(t,\varepsilon)\,x_1(t,\varepsilon) + \cdots + A_{pp}(t,\varepsilon)\,x_p(t,\varepsilon) \end{aligned}\right\}$$

$$R(\varepsilon)\,x(a,\varepsilon) + S(\varepsilon)\,x(b,\varepsilon) = c(\varepsilon)$$

as $\varepsilon\to 0^+$ and the related "degenerate" problem obtained by setting $\varepsilon=0$. Here the h_i are integers, $0<h_2<\cdots<h_p$, x_i is a vector of dimension n_i, $A_{ij}(t,\varepsilon)$ are matrices of appropriate orders with asymptotic expansions, x is the vector $\begin{pmatrix} x_1\\ \vdots\\ x_p\end{pmatrix}$, R and S are square matrices of order $\sum\limits_{i=1}^{p} n_i$ and $\varepsilon>0$.

It is shown that under certain conditions the solution of the "complete" problem as $\varepsilon\to 0^+$ approaches a solution of the "degenerate" differential system and satisfies n_1 appropriate "degenerate" boundary conditions.

1. Introduction

We are concerned with showing the relationship of the solution of a boundary problem (1.1), (1.2) as $\varepsilon \to 0^+$ to the solutions of a related degenerate problem (1.3), (1.4). The problems are

$$
\begin{aligned}
\frac{d}{dt} x_1(t,\varepsilon) &= A_{11}(t,\varepsilon)\, x_1(t,\varepsilon) + \cdots + A_{1p}(t,\varepsilon)\, x_p(t,\varepsilon)\\
\varepsilon^{h_2}\frac{d}{dt} x_2(t,\varepsilon) &= A_{21}(t,\varepsilon)\, x_1(t,\varepsilon) + \cdots + A_{2p}(t,\varepsilon)\, x_p(t,\varepsilon)\\
&\;\vdots\\
\varepsilon^{h_p}\frac{d}{dt} x_p(t,\varepsilon) &= A_{p1}(t,\varepsilon)\, x_1(t,\varepsilon) + \cdots + A_{pp}(t,\varepsilon)\, x_p(t,\varepsilon),
\end{aligned}
\tag{1.1}
$$

$$
R(\varepsilon)\, x(a,\varepsilon) + S(\varepsilon)\, x(b,\varepsilon) = c(\varepsilon), \tag{1.2}
$$

$$
\begin{aligned}
\frac{d}{dt} x_1 &= A_{11}(t,0)\, x_1 + \cdots + A_{1p}(t,0)\, x_p\\
0 &= A_{21}(t,0)\, x_1 + \cdots + A_{2p}(t,0)\, x_p\\
&\;\vdots\\
0 &= A_{p1}(t,0)\, x_1 + \cdots + A_{pp}(t,0)\, x_p,
\end{aligned}
\tag{1.3}
$$

$$
R(0)\, x(a) + S(0)\, x(b) = c(0), \tag{1.4}
$$

where the h_i are integers, $0 < h_2 < h_3 < \cdots < h_p = h$, x_i is a vector of dimension n_i, $m = \sum_{j=2}^{p} n_j$, $A(t,\varepsilon)_{ij}$ are matrices of appropriate orders with asymptotic expansions, x is the vector $\begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$, R and S are square matrices of order $(n_1 + m)$ and $\varepsilon > 0$.

Under three hypotheses, H 1, H 2, and H 3, we shall prove Theorem 1 (Section 6), which embodies our results for the problem indicated above. We begin by reducing the problem (1.1), (1.2) to a canonical form (2.15), (6.3). We show that the solution of the canonical boundary problem has a limit as $\varepsilon \to 0^+$ which satisfies the corresponding degenerate differential system and n_1 of the degenerate boundary conditions. A discussion of the hypotheses is given in Section 7.

The author wishes to express his deep appreciation to Professor H. L. Turrittin for his guidance and encouragement during the preparation of this paper.

2. Preliminary Transformations

If $A_{pp}(t,0)$ is non-singular, $a \leqq t \leqq b$, the last equation of (1.3) may be solved for x_p in terms of $x_1, \ldots, x_{p-1}$ and substituted into the preceeding equations of (1.3) to give a system of the same form as (1.3) in $x_1, \ldots, x_{p-1}$. The last equation of this form is

$$
0 = (A_{p-1,1} - A_{p-1,p} A_{pp}^{-1} A_{p1})\, x_1 + \cdots + (A_{p-1,p-1} - A_{p-1,p} A_{pp}^{-1} A_{p,p-1})\, x_{p-1}.
$$

Thus, if $(A_{p-1,p-1} - A_{p-1,p} A_{pp}^{-1} A_{p,p-1})$ is non-singular, $a \leqq t \leqq b$, we may solve this equation for x_{p-1} in terms of $x_1, \ldots, x_{p-2}$ and repeat the process until we solve the first equation for x_1, which is a differential equation. If this process can be carried out completely, step by step, and $x_p, x_{p-1}, \ldots, x_2$ all eliminated,

then E. R. RANG [8] has shown there exists a non-singular transformation, $0 \leqq \varepsilon \leqq \varepsilon_0$, $a \leqq t \leqq b$, of the form

$$x(t, \varepsilon) = \left(T_1(t) + \varepsilon\, Q_1(t)\right)\left(T_2(t) + \varepsilon^2 Q_2(t)\right) \cdots \left(T_L(t) + \varepsilon^L Q_L(t)\right) y(t, \varepsilon) \tag{2.1}$$

which changes the differential system (1.1) into the form (2.2) with corresponding degenerate differential system (2.3).

$$\begin{aligned} \frac{d}{dt} y_1(t, \varepsilon) &= C_{11}(t, \varepsilon)\, y_1(t, \varepsilon) + \cdots C_{1p}(t, \varepsilon)\, y_p(t, \varepsilon) \\ \varepsilon^{h_2} \frac{d}{dt} y_2(t, \varepsilon) &= C_{21}(t, \varepsilon)\, y_1(t, \varepsilon) + \cdots C_{2p}(t, \varepsilon)\, y_p(t, \varepsilon) \\ &\vdots \\ \varepsilon^{h_p} \frac{d}{dt} y_p(t, \varepsilon) &= C_{p1}(t, \varepsilon)\, y_1(t, \varepsilon) + \cdots C_{pp}(t, \varepsilon)\, y_p(t, \varepsilon), \end{aligned} \tag{2.2}$$

$$\begin{aligned} \frac{d}{dt} y_1 &= C_{11}(t, 0)\, y_1 + \cdots + C_{1p}(t, 0)\, y_p \\ 0 &= C_{21}(t, 0)\, y_1 + \cdots + C_{2p}(t, 0)\, y_p \\ &\vdots \\ 0 &= C_{p1}(t, 0)\, y_1 + \cdots + C_{pp}(t, 0)\, y_p \end{aligned} \tag{2.3}$$

where $C_{jj}(t, 0)$ is non-singular $a \leqq t \leqq b$, $j = 2, \ldots, p$, and in particular

$$C_{pp}(t, 0) = A_{pp}(t, 0) \quad \text{and} \quad C_{p-1, p-1} = A_{p-1, p-1} - A_{p-1, p} A_{pp}^{-1} A_{p, p-1}, \quad \text{etc.}$$

Further, L may be chosen so large that the elements of $C_{ij}(t, \varepsilon)$ are $O(\varepsilon^\alpha)$ for any particular given large integer $\alpha > 0$, $i \neq j$. We shall have occasion to assume that this has been done.

For an N^{th} order differential system of the form

$$\varepsilon^\beta \frac{dz}{dt} = \left(\sum_{j=0}^{\infty} a_j(t)\, \varepsilon^j\right) z \tag{2.4}$$

the most general asymptotic expansions of solutions have been given by H. L. TURRITTIN [9]. In particular, he has given sufficient conditions for the existence of a transformation $z = H(t, \varepsilon)\, w(t, \varepsilon)$, where

$$H(t, \varepsilon) = \sum_{k=0}^{K} \varepsilon^{k/r} H_k(t), \qquad r \text{ and } K \text{ suitable positive integers,} \tag{2.5}$$

which will transform the equation (2.4) into a new equation

$$\varepsilon^\beta \frac{dw}{dt} = \left\{(\delta_{ij} \lambda_j(t, \varepsilon)) + \varepsilon^\beta B(t, \varepsilon)\right\} w(t, \varepsilon) \tag{2.6}$$

where $B(t, \varepsilon) - O(1)$ and the characteristic polynomials $\lambda_j(t, \varepsilon)$ have the form

$$\lambda_j(t, \varepsilon) = \sum_{k=0}^{r\beta - 1} \varepsilon^{k/r} \lambda_{jk}(t), \qquad j = 1, 2, \ldots, N, \tag{2.7}$$

and $\lambda_j(t, \varepsilon) \equiv \lambda_k(t, \varepsilon)$, or $\lambda_j(t, \varepsilon) \neq \lambda_k(t, \varepsilon)$, $a \leqq t \leqq b$, $0 < \varepsilon \leqq \varepsilon_0$.

If $r > 1$, the problem could have been adjusted in the beginning by introduction of a new parameter $\varepsilon^{1/r} = \varepsilon_1$. We assume that this has been done, so that without loss of generality we may assume $r = 1$.

Further, if the characteristic polynomials $\lambda_j(t, \varepsilon)$ are such that

$$\mathrm{Re}\{\varepsilon^{-\beta}\lambda_1(t,\varepsilon)\} \leqq \cdots \leqq \mathrm{Re}\{\varepsilon^{-\beta}\lambda_N(t,\varepsilon)\}, \quad \text{for} \quad a \leqq t \leqq b \quad \text{and} \quad 0 < \varepsilon \leqq \varepsilon_0,$$

there exists a fundamental matrix solution of (2.6) of the form $W(t, \varepsilon) = F(t, \varepsilon)\, E(t, \varepsilon)$, where when $a \leqq t \leqq b$ and $0 < \varepsilon \leqq \varepsilon_0$

$$F(t,\varepsilon) = \begin{pmatrix} F_{11} & \dots & F_{1M} \\ \vdots & & \\ F_{M1} & \dots & F_{MM} \end{pmatrix}; \qquad E(t,\varepsilon) = \begin{pmatrix} E_1 & 0 & \dots & 0 \\ 0 & E_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & & E_M \end{pmatrix},$$

and $F_{ij} \sim \varepsilon^{\beta_{ij}} \sum_{k=0}^{\infty} F_{ijk}(t)\, \varepsilon^k$. Here $\beta_{ij} > 0$ if $i \neq j$; $\beta_{ii} = 0$; and $F_{ii0}(t)$ is non-singular when $a \leqq t \leqq b$. Also $E_i = I_i \exp\left\{\varepsilon^{-\beta} \int_a^t \lambda_{\tau_i}(\sigma, \varepsilon)\, d\sigma\right\}$, where I_i is an identity matrix, and λ_{τ_i} are the distinct characteristic polynomials.

Thus a transformation of the type

$$y(t,\varepsilon) = \begin{pmatrix} I_1 & 0 & \dots & 0 \\ 0 & H_2(t,\varepsilon) & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & H_p(t,\varepsilon) \end{pmatrix} z(t,\varepsilon) \tag{2.8}$$

will change (2.2) into the form

$$\begin{aligned} \frac{dz_1}{dt} &= C_{11}(t,\varepsilon)\, z_1 + \{C_{12}(t,\varepsilon)\, H_2(t,\varepsilon)\}\, z_2 + \cdots + \{C_{1p}(t,\varepsilon)\, H_p(t,\varepsilon)\}\, z_p \\ \varepsilon^{h_2} \frac{dz_2}{dt} &= \{H_2^{-1} C_{21}\}\, z_1 + H_2^{-1}\left\{C_{22} H_2 - \varepsilon^{h_2} \frac{d}{dt} H_2\right\} z_2 + \cdots + \{H_2^{-1} C_{2p} H_p\}\, z_p \\ &\;\;\vdots \\ \varepsilon^{h_p} \frac{dz_p}{dt} &= \{H_p^{-1} C_{p1}\}\, z_1 + \quad \cdots \quad + H_p^{-1}\left\{C_{pp} H_p - \varepsilon^{h_p} \frac{d}{dt} H_p\right\} z_p. \end{aligned} \tag{2.9}$$

Thus, if TURRITTIN'S results apply to the individual equations,

$$\varepsilon^{h_j} \frac{d}{dt} y_j(t,\varepsilon) = C_{jj}(t,\varepsilon)\, y_j(t,\varepsilon), \qquad j = 2, \dots, p, \tag{2.10}$$

as we shall assume, the transformations $H_j(t, \varepsilon)$ may be chosen so that

$$H_j^{-1}(t,\varepsilon)\left\{C_{jj}(t,\varepsilon)\, H_j(t,\varepsilon) - \varepsilon^{h_j} \frac{d}{dt} H_j(t,\varepsilon)\right\}$$

has the canonical form shown in (2.6). Also in transformation (2.1) L may be chosen such that the elements of $H_j^{-1}(t, \varepsilon)\, C_{ji}(t, \varepsilon)\, H_i(t, \varepsilon)$ are $O(\varepsilon^{h_i})$ as $\varepsilon \to 0^+$. We may assemble all the successive transformations that have been made into a single transformation $y(t, \varepsilon) = H(t, \varepsilon)\, z(t, \varepsilon)$, where

$$H(t,\varepsilon) = \sum_{j=0}^{J} H_j(t)\, \varepsilon^j, \qquad J \text{ a suitable large integer,} \tag{2.11}$$

in such a manner that this transformation will change (1.1) into

$$\frac{dz_1}{dt} = D_{11}(t,\varepsilon)\, z_1 + D_{12}(t,\varepsilon)\, z_2 \tag{2.12}$$
$$\varepsilon^h \frac{dz_2}{dt} = D_{21}(t,\varepsilon)\, z_1 + D_{22}(t,\varepsilon)\, z_2,$$

where

$$D_{12}(t,\varepsilon) = O(\varepsilon), \quad D_{21}(t,\varepsilon) = O(\varepsilon^h), \quad D_{22}(t,\varepsilon) = \left(\delta_{ij}\,\lambda_j(t,\varepsilon)\right) + \varepsilon^h \overline{D}_{22},$$

$\overline{D}_{22}(t,\varepsilon) = O(1)$, and $\lambda_j(t,\varepsilon)$, $j = 1, \ldots, m$, are the non-vanishing characteristic polynomials associated with the differential systems (2.10); and where moreover by hypothesis

$$\lambda_1 \equiv \lambda_2 \equiv \cdots \lambda_{\tau_1} \not\equiv \lambda_{\tau_1+1} \equiv \cdots \equiv \lambda_{\tau_2} \not\equiv \cdots \lambda_{\tau_{\gamma-1}} \not\equiv \lambda_{\tau_{\gamma-1}+1} \equiv \cdots \equiv \lambda_{\tau_\gamma}.$$

It is advantageous to make one more transformation on the system (2.12), namely

$$z_1 = P_1(t)\, u_1(t,\varepsilon), \quad z_2 = \begin{pmatrix} P_{21} & 0 & \ldots & 0 \\ 0 & P_{22} & & \vdots \\ \vdots & & \ddots & \\ 0 & \ldots & & P_{2\gamma} \end{pmatrix} u_2(t,\varepsilon) = P_2(t)\, u_2, \tag{2.13}$$

where $P_{2j}(t)$ is a square matrix whose order is the multiplicity of the characteristic polynomial $\lambda_{\tau_j}(t,\varepsilon)$. Thus (2.12) becomes

$$\frac{du_1}{dt} = P_1^{-1}(t)\left\{D_{11}(t,\varepsilon)\, P_1(t) - \frac{d}{dt} P_1(t)\right\} u_1 + \left\{P_1^{-1}(t)\, D_{12}(t,\varepsilon)\, P_2(t)\right\} u_2$$
$$\varepsilon^h \frac{du_2}{dt} = \left\{P_2^{-1}(t)\, D_{21}(t,\varepsilon)\, P_1(t)\right\} u_1 + P_2^{-1}(t)\left\{D_{22}(t,\varepsilon)\, P_2(t) - \varepsilon^h \frac{d}{dt} P_2(t)\right\} u_2.$$

Choose $P_1(t)$ such that $\frac{d}{dt} P_1(t) = D_{11}(t,0)\, P_1(t)$, $P(a) = I_1$, and $P_{2j}(t)$ such that $\frac{d}{dt} P_{2j}(t) = \left(\overline{D}_{22}(t,0)\right)_{jj} P_{2j}(t)$, $P_{2j}(a) = I_j$, where $\left(\overline{D}_{22}(t,0)\right)_{ij}$ is the ij^{th} block of the matrix $\overline{D}_{22}$ partitioned in the same manner as P_2.

Thus the transformation

$$y(t,\varepsilon) = H(t,\varepsilon) \begin{pmatrix} P_1(t) & 0 \\ 0 & P_2(t) \end{pmatrix} u(t,\varepsilon) \tag{2.14}$$

will change (1.1) into the *canonical differential system*

$$\frac{d}{dt} u_1(t,\varepsilon) = B_{11}(t,\varepsilon)\, u_1(t,\varepsilon) + B_{12}(t,\varepsilon)\, u_2(t,\varepsilon) \tag{2.15}$$
$$\varepsilon^h \frac{d}{dt} u_2(t,\varepsilon) = B_{21}(t,\varepsilon)\, u_1(t,\varepsilon) + B_{22}(t,\varepsilon)\, u_2(t,\varepsilon)$$

with a related *canonical degenerate differential system*

$$\frac{du_1}{dt} = 0 \tag{2.16}$$
$$0 = \left(\delta_{ij}\,\alpha_j(t)\right) u_2$$

where $\alpha_i(t)$ are the characteristic roots of the matrices $C_{ii}(t,0)$ of the equation (2.2), all of which were assumed to be non-zero on $a \leqq t \leqq b$, which implies in

turn that the solution of the canonical degenerate differential system is $u_1 = c_1$, $u_2 = 0$, where c_1 is an arbitrary constant vector of appropriate order.

Further, this conversion of the differential system (1.1) into (2.15) is such that a fundamental matrix solution $W(t, \varepsilon)$ for the system (2.15), when TURRITTIN's results apply, has the form

$$W(t, \varepsilon) = ([I])\, E(t, \varepsilon)^{\star}. \tag{2.17}$$

Since we are treating the boundary problem as a whole, the transformation (2.14) will affect the boundary form (1.2) as well. This effect will be considered in more detail in Section 4.

3. The Canonical Problem

We make the following hypothesis.

H 1: (i) The matrices $A_{ij}(t, \varepsilon)$ *indicated in* (1.1) *have asymptotic expansions of appropriate high finite orders.*

(ii) The matrices $A_{pp}(t, 0)$, $A_{p-1, p-1}(t, 0) - A_{p-1, p}(t, 0)\, A_{pp}^{-1}(t, 0)\, A_{p, p-1}(t, 0)$ *and similar matrices referred to in Section* 2 *are non-singular on the interval* $a \leqq t \leqq b$.

(iii) There exists a non-singular transformation $x(t, \varepsilon) = \overline{H}(t, \varepsilon)\, u(t, \varepsilon)$, *where* $\overline{H}(t, \varepsilon) = H_0(t) + \varepsilon H_1(t) + \cdots \varepsilon^J H_J(t)$, $a \leqq t \leqq b$, $0 \leqq \varepsilon \leqq \varepsilon_0$, *which will convert* (1.1) *into the canonical form* (2.15).

(iv) The m non-vanishing characteristic polynomials $\lambda_j(t, \varepsilon)$ *satisfy*

$$\mathrm{Re}\{\varepsilon^{-h} \lambda_1(t, \varepsilon)\} \leqq Re\{\varepsilon^{-h} \lambda_2(t, \varepsilon)\} \leqq \cdots \leqq \mathrm{Re}\{\varepsilon^{-h} \lambda_m(t, \varepsilon)\},$$
$$a \leqq t \leqq b, \qquad 0 < \varepsilon \leqq \varepsilon_0.$$

If $X(t, \varepsilon)$ is any fundamental matrix for a system of differential equations of the form $\frac{dx}{dt} = A(t, \varepsilon)\, x(t, \varepsilon)$, $\varepsilon > 0$, $a \leqq t \leqq b$, then any particular vector solution $l(t, \varepsilon)$ must be of the form $l(t, \varepsilon) = X(t, \varepsilon)\, l(\varepsilon)$. Thus, if $l(t, \varepsilon)$ is to satisfy the boundary conditions $R(\varepsilon)\, l(a, \varepsilon) + S(\varepsilon)\, l(b, \varepsilon) = c(\varepsilon)$, we must have $\{R(\varepsilon)\, X(a, \varepsilon) + S(\varepsilon)\, X(b, \varepsilon)\}\, l(\varepsilon) = c(\varepsilon)$. Thus, if $\Delta(\varepsilon) = \{R(\varepsilon)\, X(a, \varepsilon) + S(\varepsilon)\, X(b, \varepsilon)\}$, $l(t, \varepsilon)$ will be unique if $\Delta(\varepsilon)$ is non-singular, for in this event

$$l(t, \varepsilon) = X(t, \varepsilon)\, \Delta^{-1}(\varepsilon)\, c(\varepsilon). \tag{3.1}$$

The limit problem is then the computation of the

$$\lim_{\varepsilon \to 0^+} l(t, \varepsilon) = \lim_{\varepsilon \to 0^+} X(t, \varepsilon)\, \Delta^{-1}(\varepsilon)\, c(\varepsilon).$$

To evaluate this limit we need more detailed information about the structure of $\Delta^{-1}(\varepsilon)$.

4. $\Delta^{-1}(\epsilon)$ for a Canonical Problem

We assume that we are dealing with the canonical differential system (2.15) which has been obtained from (1.1) by the transformation $x(t, \varepsilon) = \overline{H}(t, \varepsilon)\, u(t, \varepsilon)$ of H 1—(iii). This transformation will change the boundary conditions from (1.2) into

$$M(\varepsilon)\, u(a, \varepsilon) + N(\varepsilon)\, u(b, \varepsilon) = c(\varepsilon) \tag{4.1}$$

$\star$ $[\varphi(t)]$ represents a function $\varphi(t, \varepsilon) = \varphi(t) + \varphi_1(t, \varepsilon)\, \varepsilon^{\gamma}$, $\gamma > 0$, $|\varphi_1(t, \varepsilon)| < B$.

where

$$M(\varepsilon) = R(\varepsilon)\,\overline{H}(a,\varepsilon) \quad \text{and} \quad N(\varepsilon) = S(\varepsilon)\,\overline{H}(b,\varepsilon).$$

We make the following hypothesis.

H 2: (i) $R(\varepsilon) = R_0 + \varepsilon R_1(\varepsilon)$, $S(\varepsilon) = S_0 + \varepsilon S_1(\varepsilon)$ *where the elements of* $R_1(\varepsilon)$ *and* $S_1(\varepsilon)$ *are bounded for* $0 \leqq \varepsilon \leqq \varepsilon_0$ *and the rank of* $(R(\varepsilon) : S(\varepsilon)) = n_1 + m$ *for* $0 \leqq \varepsilon \leqq \varepsilon_0$.

(ii) The non-vanishing characteristic polynomials have non-zero real parts, i.e.,

$$\mathrm{Re}\{\varepsilon^{-h}\lambda_1(t,\varepsilon)\} \leqq \cdots \leqq \mathrm{Re}\{\varepsilon^{-h}\lambda_k(t,\varepsilon)\} < 0 < \mathrm{Re}\{\varepsilon^{-h}\lambda_{k+1}(t,\varepsilon)\} \leqq \cdots \leqq \mathrm{Re}\{\varepsilon^{-h}\lambda_m\}$$

$$\text{on} \quad a \leqq t \leqq b, \qquad 0 < \varepsilon \leqq \varepsilon_0.$$

If we choose for the fundamental matrix the one indicated in (2.17), we have

$$\Delta(\varepsilon) = \left\{ M(\varepsilon)\,([I]) + N(\varepsilon)\,([I]) \begin{pmatrix} I_{n_1} & 0 \\ 0 & E_m(b,\varepsilon) \end{pmatrix} \right\}.$$

If $D(\varepsilon) = \det \Delta(\varepsilon) \neq 0$ $0 < \varepsilon \leqq \varepsilon_0$, $\Delta(\varepsilon)$ will be non-singular. We have

$$D(\varepsilon) \sim \Sigma_\alpha\, a_\alpha(\varepsilon)\, e^{\omega_\alpha(\varepsilon)},$$

where

(i) α covers some finite range,

(ii) $\omega_\alpha(\varepsilon)$ are distinct quantities, each of which is of the form $\sum_{j=I}^{J} \varrho_{k_j}(b,\varepsilon)$, where $I, J = 0, 1, \ldots, m$; $J \geqq I$ and $(k_0, k_1, \ldots, k_m)$ is any permutation of $(0, 1, \ldots, m)$; $\varrho_0(b,\varepsilon) \equiv 0$, $\varrho_j(b,\varepsilon) = \varepsilon^{-h} \int_a^b \lambda_j(\sigma,\varepsilon)\, d\sigma$, $j \geqq 1$,

(iii) the coefficient functions $a_\alpha(\varepsilon) \not\equiv 0$, $a_\alpha(\varepsilon) = O(1)$ as $\varepsilon \to 0^+$.

For a discussion of the zeros of such exponential sums, see TURRITTIN [*10*].

All terms indicated in (ii) need not be present; however, one is of particular interest, namely the term $a(\varepsilon)\, e^{\omega(\varepsilon)}$ where $\omega(\varepsilon) = \sum_{j=k+1}^{m} \varrho_j(b,\varepsilon)$. (We note that, if $k = m$, then $\omega(\varepsilon) \equiv 0$.) An explicit expression for the leading term of the coefficient function $a(\varepsilon)$ can be given as follows. Let the columns of $R(0)\,\overline{H}(a,\varepsilon)$ and $S(0)\,\overline{H}(b,\varepsilon)$ be α_{i1}, α_{i2} respectively, *i.e.*, $R(0)\,\overline{H}(a,\varepsilon) = (\alpha_{11} : \cdots : \alpha_{n_1+m,1})$ $S(0)\,\overline{H}(b,\varepsilon) = (\alpha_{12} : \cdots \alpha_{n_1+m,2})$ and let $\Omega(\varepsilon)$ be the $n_1 + m^{\text{th}}$ order square matrix

$$\Omega = (\alpha_{11} + \alpha_{12} : \cdots : \alpha_{n_1,1} + \alpha_{n_1,2} : \alpha_{n_1+1,1} : \alpha_{n_1+2,1} : \cdots \\ \cdots : \alpha_{n_1+k,1} : \alpha_{n_1+k+1,2} : \cdots : \alpha_{n_1+m,2}). \tag{4.2}$$

The leading term of $a(\varepsilon)$ is the determinant of $\Omega(\varepsilon)$. In general for ε sufficiently small it can be shown that if $\Omega(\varepsilon)$ is non-singular (see for instance HARRIS [*5*], page 88),

$$\Delta(\varepsilon) = ([\Omega]) \begin{pmatrix} I_{n_1} & 0 & 0 \\ 0 & I_k & 0 \\ 0 & 0 & E_{m-k}(b,\varepsilon) \end{pmatrix},$$

where

$$E_{m-k}(b,\varepsilon)=\left(\delta_{ij}\exp\left\{\varepsilon^{-h}\int_a^b\lambda_{k+j}(t,\varepsilon)\,dt\right\}\right)$$
$$=\left(\delta_{ij}\exp\{\varrho_{k+j}(b,\varepsilon)\}\right),\qquad j=1,2,\ldots,m-k,$$

and hence

$$\text{(4.3)}\qquad \Delta^{-1}(\varepsilon)=\begin{pmatrix} I_{n_1} & 0 & 0\\ 0 & I_k & 0\\ 0 & 0 & E^{-1}_{m-k}(b,\varepsilon)\end{pmatrix}\left([\Omega^{-1}(\varepsilon)]\right).$$

It will now be shown that the computation of the limit of (3.1) as $\varepsilon\to 0^+$, *i.e.*, the evaluation of $\lim\limits_{\varepsilon\to 0^+} X(t,\varepsilon)\,\Delta^{-1}(\varepsilon)\,c(\varepsilon)$, is essentially the question of the limit as $\varepsilon\to 0^+$ of the first n_1 rows of the matrix $\Omega^{-1}(\varepsilon)$. To establish the nature of the elements in the first n_1 rows of $\Omega^{-1}(\varepsilon)$ we must examine the transformation $\overline{H}(t,\varepsilon)$ of Hypothesis H 1—(iii). $\overline{H}(t,\varepsilon)$ is a product of four transformations, (2.1), (2.8), (2.13) and $\begin{pmatrix} I & 0\\ 0 & P_0\end{pmatrix}$, where P_0 is a constant matrix which renumbered the last m unknowns so that the characteristic polynomials are ordered by their real parts as shown in H 1—(iv).

$$\overline{H}(t,\varepsilon)=\left\{\prod_{j=1}^{L}\left(T_j(t)+\varepsilon^j Q_j(t)\right)\right\}\begin{pmatrix} I & 0 & \ldots & 0\\ 0 & H_2 & & \vdots\\ \vdots & & \ddots & \vdots\\ 0 & \ldots & & H_p\end{pmatrix}\begin{pmatrix} P_1 & 0\\ 0 & P_2\end{pmatrix}\begin{pmatrix} I & 0\\ 0 & P_0\end{pmatrix},$$

where the multiplication indicated by $\prod$ is a polynomial in ε with a non-singular matrix for coefficient of ε^0. This transformation does not change the essential character of the boundary form $R(\varepsilon)\,x(a,\varepsilon)+S(\varepsilon)\,x(b,\varepsilon)=c(\varepsilon)$, since

$$\overline{R}(\varepsilon)=R(\varepsilon)\left(\prod_{j=1}^{L}T_j(a)+\varepsilon^j Q_j(a)\right)=\left([R_0\,T(a)]\right)=\left([\overline{R}_0]\right),$$
$$\overline{S}(\varepsilon)=S(\varepsilon)\left(\prod_{j=1}^{L}T_j(b)+\varepsilon^j Q_j(b)\right)=\left([S_0\,T(b)]\right)=\left([\overline{S}_0]\right),$$

where

$$T(t)=T_1\,T_2\ldots T_L.$$

Let us consider the effect of the next transformation on $\overline{R}(\varepsilon)$. Partition $\overline{R}(\varepsilon)$ into blocks of columns containing $n_1, n_2, \ldots, n_p$, columns respectively; *i.e.*, $\overline{R}(\varepsilon)=(\overline{R}_1:\overline{R}_2:\cdots:\overline{R}_p)$. Thus

$$\overline{R}(\varepsilon)\begin{pmatrix} I & 0 & \ldots & 0\\ 0 & H_2 & & \vdots\\ \vdots & & \ddots & 0\\ 0 & \ldots & 0 & H_p\end{pmatrix}=\left(\overline{R}_1:\overline{R}_2\,H_2(a,\varepsilon):\cdots:\overline{R}_p\,H_p(a,\varepsilon)\right),$$

and the matrix $H_i(a,\varepsilon)$ affects only the matrix $\overline{R}_i(\varepsilon)$. Let us consider one such product $\overline{R}(\varepsilon)\,H(a,\varepsilon)$, momentarily dropping the subscript i. $H(a,\varepsilon)$ is a finite product of two types of transformations $\mathscr{L}_j(a,\varepsilon)$ and $\mathscr{M}_j(\varepsilon)$; more precisely (see

TURRITTIN [9], page 96),

$$H(t, \varepsilon) = \mathscr{L}_1(t, \varepsilon)\, \mathscr{M}_1(\varepsilon)\, \mathscr{L}_2(t, \varepsilon)\, \mathscr{M}_2(\varepsilon) \ldots \mathscr{L}_\beta(t, \varepsilon)\, \mathscr{M}_\beta(\varepsilon),$$

$$\mathscr{L}_j(t, \varepsilon) = L_{j0}(t) + \varepsilon L_{j1}(t) + \cdots \varepsilon^{\beta_j} L_{j\beta_j}(t), \quad L_{j0}(t) \text{ non-singular on } \quad a \leqq t \leqq b,$$

and

$$\mathscr{M}_j(\varepsilon) = (\delta_{kl}\, \varepsilon^{(k_{j_2}-k)\mu_{jk}}),$$

where n is the order of $H(t, \varepsilon)$, k_{j_1} and k_{j_2} positive integers, and

$$\mu_{jk} = \begin{cases} 1, & \text{if } 0 < k_{j_1} \leqq k \leqq k_{j_2} \leqq n \\ 0, & \text{otherwise.} \end{cases}$$

For example, a typical $\mathscr{M}(\varepsilon)$ would be

$$\begin{pmatrix} 1 & 0 & & & & \cdots & & & & 0 \\ 0 & \ddots & & & & & & & & \\ & & 1 & 0 & & & & & & \\ \cdot & & 0 & \varepsilon^l & & & & & & \\ & & & & \varepsilon^{l-1} & & & & & \\ \cdot & & & & & \ddots & & & & \\ & & & & & & \varepsilon & 0 & & \\ \cdot & & & & & & 0 & 1 & & \\ & & & & & & & & \ddots & 0 \\ 0 & & & & \cdots & & & & 0 & 1 \end{pmatrix}.$$

Consider $\bar{R}(\varepsilon)\, \mathscr{L}_1(a, \varepsilon)$. Since $\mathscr{L}_1(a, \varepsilon)$ and $\bar{R}(\varepsilon)$ have the same essential features, so does $\bar{R}(\varepsilon)\, \mathscr{L}_1(a, \varepsilon)$, *i.e.* $\bar{R}(\varepsilon)\, \mathscr{L}_1(a, \varepsilon) = ([\bar{R}(0)\, L_{10}(a)])$. Multiplication of $\bar{R}(\varepsilon)\, \mathscr{L}_1(a, \varepsilon)$ on the right by $\mathscr{M}_1(\varepsilon)$ multiplies some columns at the beginning and end by unity and every other column by a different power of ε. For definiteness, let the highest power of ε be l. Multiplication of $\bar{R}(\varepsilon)\, \mathscr{L}_1(a, \varepsilon)\, \mathscr{M}_1(\varepsilon)$ on the right by $\mathscr{L}_a(a, \varepsilon)$ replaces each column by a linear combination of the columns of $\bar{R}\, \mathscr{L}_1\, \mathscr{M}_1$, and each column of $\bar{R}\, \mathscr{L}_1\, \mathscr{M}_1$ has ε^0 for the lowest possible power of ε in the columns. Thus, if $n_1 + m - n$ columns of constants were adjoined to the matrix $\bar{R}(\varepsilon)\, \mathscr{L}_1(a, \varepsilon)\, \mathscr{M}_1(\varepsilon)\, \mathscr{L}_2(a, \varepsilon)$ and the determinant of this augmented matrix were evaluated, the lowest power of ε that could occur in the expansion of this determinant would be $\varepsilon^{\frac{l(l+1)}{2}}$ and *not* ε^0 due to certain column dependence of the matrix $\bar{R}(\varepsilon)\, \mathscr{L}_1(a, \varepsilon)\, \mathscr{M}_1(\varepsilon)\, \mathscr{L}_2(a, \varepsilon)$. Multiplication by the remaining $\mathscr{L}$ and $\mathscr{M}$ transformations and a similar argument gives rise to the following statement. Let

$$\det H_i(t, \varepsilon) = \varepsilon^{q_i}[h_i(t)], \qquad h_i(t) \neq 0, \quad a \leqq t \leqq b. \tag{4.4}$$

Any matrix with $n_1 + m - n_i$ columns of constants adjoined to the n_i columns of $\bar{R}_i(\varepsilon)\, H_i(a, \varepsilon)$, $i = 2, \ldots, p$, will have a determinant which, when expanded in powers of ε, will begin with the term in ε^{q_i} or possibly some higher power of ε. Similar results apply to the matrix $\bar{S}_i(\varepsilon)\, H_i(b, \varepsilon)$, $i = 2, \ldots, p$. Further, if we select any k_i columns from $\bar{R}_i(\varepsilon)\, H_i(a, \varepsilon)$, say $a_1, a_2, \ldots, a_{k_i}$, and $n_i - k_i$ columns

from $\overline{S}_i(\varepsilon)\, H_i(b, \varepsilon)$, say $a_{k_i+1}, \ldots, a_{n_i}$, such that $(a_i, \ldots, a_{n_i})$ is any permutation of $(1, 2, \ldots, n_i)$ and adjoin $n_1 + m - n_i$ columns of constants, the resulting matrix often has a determinant for which the lowest order of ε that occurs is ε^{q_i}. The last two transformations are of the type already considered and do not change the previous statement.

Let us partition M, N, and Ω in the following manner:

$$M(\varepsilon) = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}, \qquad N(\varepsilon) = \begin{pmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{pmatrix}, \qquad \Omega = \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix},$$

where M_{11}, N_{11} and Ω_{11} have order n_1 and M_{22}, N_{22}, Ω_{22} have order m. We see that

$$\begin{aligned} M_{11}(0) + N_{11}(0) &= \Omega_{11}(0) \\ M_{21}(0) + N_{21}(0) &= \Omega_{21}(0), \end{aligned} \tag{4.5}$$

i.e., the elements in the first n_1 columns of Ω are $O(1)$ as $\varepsilon \to 0^+$. Thus, often $\det\, \Omega(\varepsilon) = O(\varepsilon^q)$ where $q = \sum_{2}^{p} q_i$, (see 4.4), and the cofactor of any element in the first n_1 columns of Ω is $O(\varepsilon^q)$.

We make the hypothesis,

H 3: The matrix $\Omega(\varepsilon)$ as given in (4.2) *satisfies the conditions:*

(i) $\Omega(\varepsilon)$ is non-singular, $0 < \varepsilon \leqq \varepsilon_0$,

(ii) the elements in the first n_1 rows of $\Omega^{-1}(\varepsilon)$ are $O(1)$ as $\varepsilon \to 0^+$.

Hypothesis H 3—(i) assures us that $\Delta^{-1}(\varepsilon)$ has the form shown in (4.3).

5. Evaluation of $\lim\limits_{\epsilon \to 0^+} X(t, \epsilon)\, \Delta^{-1}(\epsilon)\, c(\epsilon)$

Combining (2.17) and (4.3), we have the following representation of the unique solution of (2.15) and (4.1):

$$\begin{aligned} l(t, \varepsilon) &= X(t, \varepsilon)\, \Delta^{-1}(\varepsilon)\, c(\varepsilon) \\ &= ([I]) \begin{pmatrix} I_{n_1} & 0 & 0 \\ 0 & E_k(t, \varepsilon) & 0 \\ 0 & 0 & E_{m-k}(t, \varepsilon) \end{pmatrix} \begin{pmatrix} I_{n_1} & 0 & 0 \\ 0 & I_k & 0 \\ 0 & 0 & E^{-1}_{m-k}(b, \varepsilon) \end{pmatrix} ([\Omega^{-1}(\varepsilon)])\, c(\varepsilon) \\ &= ([I]) \begin{pmatrix} I_{n_1} & 0 & 0 \\ 0 & E_k(t, \varepsilon) & 0 \\ 0 & 0 & E_{m-k}(t, \varepsilon)\, E^{-1}_{m-k}(b, \varepsilon) \end{pmatrix} ([\Omega^{-1}(\varepsilon)])\, c(\varepsilon). \end{aligned}$$

We have

$$E_k(t, \varepsilon) = \left(\delta_{ij} \exp\left\{\varepsilon^{-h} \int_a^t \lambda_i(\sigma, \varepsilon)\, d\sigma\right\}\right), \qquad i = 1, 2, \ldots, k,$$

and

$$E_{m-k}(t, \varepsilon)\, E^{-1}_{m-k}(b, \varepsilon) = \left(\delta_{ij} \exp\left\{-\,\varepsilon^{-h} \int_t^b \lambda_{i+k}(\sigma, \varepsilon)\, d\sigma\right\}\right), \qquad i = 1, 2, \ldots, m-k,$$

and hence*

$$E_k(t,\varepsilon) \to O_k, \qquad a < \delta_1 \leqq t \leqq b,$$

$$E_{m-k}(t,\varepsilon)\, E^{-1}_{m-k}(b,\varepsilon) \to O_{m-k}, \qquad a \leqq t \leqq \delta_2 < b,$$

exponentially fast due to H 2—(ii), and uniformly in t for the indicated intervals. Thus

$$(5.1)\quad \lim_{\varepsilon\to 0^+} X(t,\varepsilon)\, \Delta^{-1}(\varepsilon)\, c(\varepsilon) = \begin{pmatrix} \bar{\Omega}_{11}(0) & \bar{\Omega}_{12}(0) \\ 0 & 0 \end{pmatrix} \begin{pmatrix} c_1(0) \\ c_2(0) \end{pmatrix} = \begin{pmatrix} l_1 \\ 0 \end{pmatrix} \quad a<\delta_1 \leqq t \leqq \delta_2 < b,$$

where the vector $l_1 = \bar{\Omega}_{11}(0)\, c_1(0) + \bar{\Omega}_{12}(0)\, c_2(0)$, and

$$\Omega^{-1}(\varepsilon) = \begin{pmatrix} \bar{\Omega}_{11}(\varepsilon) & \bar{\Omega}_{12}(\varepsilon) \\ \bar{\Omega}_{21}(\varepsilon) & \bar{\Omega}_{22}(\varepsilon) \end{pmatrix}, \qquad \Omega_{11}(\varepsilon) \text{ of order } n_1.$$

(We note that, if $k=m$, b may be included, and if $k=0$, a may be included.)

It is clear that the limiting constant vector l

$$(5.2)\qquad l = \lim_{\substack{\varepsilon\to 0^+ \\ a<t<b}} l(t,\varepsilon) = \begin{pmatrix} \bar{\Omega}_{11}(0)\, c_1(0) + \bar{\Omega}_{12}(0)\, c_2(0) \\ 0 \end{pmatrix} = \begin{pmatrix} l_1 \\ 0 \end{pmatrix}$$

is defined for the interval $a \leqq t \leqq b$ and as a function of t is a solution of the degenerate differential system (2.16). We shall now show that this limiting solution l satisfies n_1 degenerate boundary conditions.

6. Boundary Conditions Satisfied by the Limiting Solution

Multiplication of the boundary form (4.1) on the left by any non-singular matrix of constants will give rise to an equivalent boundary form. We have partitioned $\Omega^{-1}(\varepsilon)$ and $\Omega(\varepsilon)$ as follows:

$$\Omega^{-1}(\varepsilon) = \begin{pmatrix} \bar{\Omega}_{11} & \bar{\Omega}_{12} \\ \bar{\Omega}_{21} & \bar{\Omega}_{21} \end{pmatrix}, \qquad \Omega(\varepsilon) = \begin{pmatrix} \Omega_{11} & \Omega_{21} \\ \Omega_{21} & \Omega_{22} \end{pmatrix};$$

hence

$$\bar{\Omega}_{11}(\varepsilon)\, \Omega_{11}(\varepsilon) + \bar{\Omega}_{12}(\varepsilon)\, \Omega_{21}(\varepsilon) = I_{n_1},$$

and the nature of $\Omega_{11}(0)$, $\Omega_{21}(0)$, and H 3—(ii) imply

$$\bar{\Omega}_{11}(0)\, \Omega_{11}(0) + \bar{\Omega}_{12}(0)\, \Omega_{21}(0) = I_{n_1},$$

and the matrix $(\bar{\Omega}_{11}(0) : \bar{\Omega}_{12}(0))$ has rank n_1. Using (4.5), we have

$$(6.1)\qquad \bar{\Omega}_{11}(0)\, \big(M_{11}(0) + N_{11}(0)\big) + \bar{\Omega}_{12}(0)\, \big(M_{21}(0) + N_{21}(0)\big) = I_{n_1}.$$

Further, there exists constant matrices F_{21} and F_{22} such that the matrix

$$(6.2)\qquad F = \begin{pmatrix} \bar{\Omega}_{11}(0) & \bar{\Omega}_{12}(0) \\ F_{21} & F_{22} \end{pmatrix}$$

is non-singular.

* The symbol O_k here represents the zero square matrix of order k.

Let us replace (4.1) by the equivalent boundary form

$$\overline{M}(\varepsilon)\, u(a, \varepsilon) + \overline{N}(\varepsilon)\, u(b, \varepsilon) = \bar{c}(\varepsilon), \tag{6.3}$$

where $\overline{M}(\varepsilon) = FM(\varepsilon)$, $\overline{N}(\varepsilon) = FN(\varepsilon)$, and $\bar{c}(\varepsilon) = Fc(\varepsilon)$. The degenerate boundary form corresponding to (6.3) is

$$\overline{M}(0)\, u(a) + \overline{N}(0)\, u(b) = \bar{c}(0), \tag{6.4}$$

where

$$\overline{M}(0) = \begin{pmatrix} \overline{M}_{11}(0) & \overline{M}_{12}(0) \\ \overline{M}_{21}(0) & \overline{M}_{22}(0) \end{pmatrix}, \quad \text{and} \quad \overline{N}(0) = \begin{pmatrix} \overline{N}_{11}(0) & \overline{N}_{12}(0) \\ \overline{N}_{21}(0) & \overline{N}_{22}(0) \end{pmatrix}.$$

Substituting the limiting solution l into the left-hand side of (6.4), we get

$$\overline{M}(0) \begin{pmatrix} l_1 \\ 0 \end{pmatrix} + \overline{N}(0) \begin{pmatrix} l_1 \\ 0 \end{pmatrix} = \begin{pmatrix} \{\overline{M}_{11}(0) + \overline{N}_{11}(0)\}\, l_1 \\ \{\overline{M}_{21}(0) + \overline{N}_{21}(0)\}\, l_1 \end{pmatrix}.$$

By direct computation we have

$$\overline{M}_{11}(0) = \overline{\Omega}_{11}(0)\, M_{11}(0) + \overline{\Omega}_{12}(0)\, M_{21}(0),$$
$$\overline{N}_{11}(0) = \overline{\Omega}_{11}(0)\, N_{11}(0) + \overline{\Omega}_{12}(0)\, N_{21}(0),$$

and hence by (6.1)

$$\overline{M}_{11}(0) + \overline{N}_{11}(0) = I_{n_1}.$$

Further,

$$\bar{c}(0) = \begin{pmatrix} \bar{c}_1(0) \\ \bar{c}_2(0) \end{pmatrix} = \begin{pmatrix} \overline{\Omega}_{11}(0) & \overline{\Omega}_{12}(0) \\ F_{21} & F_{22} \end{pmatrix} \begin{pmatrix} c_1(0) \\ c_2(0) \end{pmatrix},$$

so

$$\bar{c}_1(0) = \overline{\Omega}_{11}(0)\, c_1(0) + \overline{\Omega}_{12}(0)\, c_2(0) = l_1.$$

Thus, the limiting solution l satisfies the first n_1 degenerate boundary conditions of (6.4) corresponding to the boundary form (6.3).

Without loss of generality we may assume that the boundary form (1.2) has been replaced by the equivalent one obtained by multiplication on the left by F as given in (6.2). Further, the solution of the canonical problem will provide the solution of the original problem through the transformation (2.14).

$$y(t, \varepsilon) = H(t, \varepsilon) \begin{pmatrix} P_1(t) & 0 \\ 0 & P_2(t)\, P_0 \end{pmatrix} l(t, \varepsilon). \tag{6.5}$$

The transformation $H(t, \varepsilon)$ was defined for $a \leqq t \leqq b$, $0 \leqq \varepsilon \leqq \varepsilon_0$, and $\lim_{\varepsilon \to 0^+} H(t, \varepsilon) = H(t, 0)$ exists, $a \leqq t \leqq b$. Thus the limiting solution for the problem (1.1), (1.2) will be

$$y(t) = H(t, 0) \begin{pmatrix} P_1(t) & 0 \\ 0 & P_2(t)\, P_0 \end{pmatrix} l, \qquad a < t < b. \tag{6.6}$$

Theorem 1. *Under hypotheses H 1, H 2, H 3, the two-point boundary problem (1.1), (1.2) has a unique solution $y(t, \varepsilon)$ on $a \leqq t \leqq b$, $0 < \varepsilon \leqq \varepsilon_0$, such that the $\lim_{\varepsilon \to 0^+} y(t, \varepsilon) = y(t)$ exists on the open interval $a < t < b$, and uniformly on any closed*

sub-interval $a \leqq \delta_1 \leqq t \leqq \delta_2 < b$. The function $y(t)$ satisfies the degenerate differential system (1.3). *The limits, $y(a+0)$ and $y(b-0)$ exist and satisfy the first n_1 boundary conditions of the degenerate boundary form* (1.4).

7. Remarks

(1) H 1 is a condition of regularity imposed on the differential system to ensure that solutions to the degenerate differential system exist and that asymptotic solutions exist for small ε and t in the interval $a \leqq t \leqq b$. H 2 is a regularity condition on the boundary form and a restriction on the characteristic polynomials $\lambda_j(t, \varepsilon)$ which allows us to determine the essential character of the matrix $\Delta(\varepsilon)$. H 3 defines a "regular" problem and assures the existence of the limiting solution. H 3 could be replaced by the assumption that the coefficient function $a(\varepsilon)$ for which $\det \Omega(\varepsilon)$ is the leading term is $\neq 0$. $a(\varepsilon)$ is the determinant of some matrix $\mathscr{A}(\varepsilon)$. If $\mathscr{A}(\varepsilon)$ is non-singular, $0 < \varepsilon \leqq \varepsilon_0$, $\Delta(\varepsilon)$ has essentially the same form as before, and the limiting solution will exist only if the first n_1 rows of $\mathscr{A}^{-1}(\varepsilon)$ are $O(1)$ as $\varepsilon \to 0^+$.

Elementary considerations show that Ω is singular for the following example, but the limiting solution exists, satisfies the degenerate differential system and the first boundary condition:

$$\left\{\begin{aligned} \frac{d x_1}{dt} &- x_2(t, \varepsilon) \\ \varepsilon \frac{d x_2}{dt} &= \varepsilon x_1(t, \varepsilon) + x_2(t, \varepsilon) \end{aligned}\right\}$$

and

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} x(0, \varepsilon) + \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} x(1, 0) = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}.$$

Here we have

$$\Omega = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \qquad a(\varepsilon) = \varepsilon[1],$$

$$\lim_{\varepsilon \to 0^+} X(t, \varepsilon)\, \Delta^{-1}(\varepsilon)\, c = \begin{pmatrix} c_1 \\ 0 \end{pmatrix}, \qquad 0 \leqq t \leqq \delta < 1,$$

and

$$(1\ 0)\begin{pmatrix} c_1 \\ 0 \end{pmatrix} + (0\ 0)\begin{pmatrix} c_1 \\ 0 \end{pmatrix} - \begin{pmatrix} c_1 \\ 0 \end{pmatrix}.$$

However, if we consider the same differential system with the boundary conditions

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} x(0, \varepsilon) + \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} x(1, \varepsilon) = c,$$

$$\Omega = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \qquad a(\varepsilon) = \varepsilon[1],$$

and we see that the $\lim_{\varepsilon \to 0^+} X(t, \varepsilon)\, \Delta^{-1}(\varepsilon)\, c$ does not exist.

(2) The results of this paper are most closely related to the work of G. G. CHAPIN Jr. [*2*], W. R. WASOW [*11*], and I. S. GRADSTEIN [*4*], who consider single N^{th} order differential equations. CHAPIN and WASOW consider the case where α

conditions are specified at one point and $N-\alpha$ at another point. CHAPIN'S differential equation is more general than that of WASOW, but his assumptions on the characteristic roots (when his problem is restricted to that of WASOW) are more stringent than WASOW'S. GRADSTEIN considers an initial value problem which is essentially contained in Theorem 1 of this paper. For our treatment of the initial value problem, we set $R \equiv I$, $S=0$, and the requirement that Ω be non-singular (or $\mathscr{A}(\varepsilon)$) implies that all the characteristic polynomials have negative real parts. A similar treatment of this type of initial value problem for non-linear systems is given by J. J. LEVIN & N. LEVINSON [7] without recourse to asymptotic expansions.

This research was supported in part by the United States Army, contract No. DA-11-022-ORD-2042.

References

[1] BIRKHOFF, G. D., & R. E. LANGER: The boundary problem and developments associated with a system of ordinary differential equations of first order. Proc. Amer. Acad. Arts Sci. **58** (2), 51—128 (1923).

[2] CHAPIN, G. G., Jr.: One and two point boundary value problems for ordinary linear differential equations containing a parameter. Ph. D. thesis, University of Minnesota, 1959.

[3] CODDINGTON, E., & N. LEVINSON: A boundary value problem for a non-linear differential equation with a small parameter. Proc. Amer. Math. Soc. **3**, 73—81 (1952).

[4] GRADSTEIN, I. S.: Linear equations with small parameters. Math. Sbornik N. S. 27 [Russian], **69**, 47—68 (1950). Also Translation Number 82, Amer. Math. Soc. 1953.

[5] HARRIS, W. A., Jr.: A boundary value problem for a system of ordinary linear differential equations involving powers of a parameter. Technical Report 3, OOR Project Number 1469, Department of Math., Univ. of Minn. 1958, 137 p.

[6] HARRIS, W. A., Jr., & H. L. TURRITTIN: Simplification of systems of linear differential equations involving a turning point. Technical Report 2, OOR Project Number 1469, Department of Math., Univ. of Minn. 1957, 35 p.

[7] LEVIN, J. J., & N. LEVINSON: Singular perturbations of nonlinear systems of differential equations and associated boundary layer equations. J. Rational Mech. Anal. **3**, 247—270 (1954).

[8] RANG, E. R.: Periodic solutions of singular perturbation problems. Ph. D. thesis, University of Minnesota, 1957.

[9] TURRITTIN, H. L.: Asymptotic expansions of solutions of systems of ordinary linear differential equations containing a parameter. Annals of Math. Studies **29**, 81—116 (1952).

[10] TURRITTIN, H. L.: Asymptotic distribution of zeros for certain exponential sums. Amer. J. Math. **66**, 199—228 (1944).

[11] WASOW, W. R.: On the asymptotic solution of a boundary value problem for ordinary differential equations containing a parameter. J. Math. Phys. **23**, 173—183 (1946).

University of Minnesota
Minneapolis

(Received January 7, 1960)

Uniqueness of Radiative Solutions to the Schroedinger Wave Equation

C. ZEMACH & F. ODEH

Communicated by W. MAGNUS

1. Introduction

This paper is primarily concerned with the uniqueness question for the exterior Dirichlet or Neumann problem associated with the Schroedinger wave equation with a suitably restricted real potential $V(\boldsymbol{r})$:

$$[\nabla^2 + k^2 - V(\boldsymbol{r})]\,\psi(\boldsymbol{r}) = 0, \qquad k > 0. \tag{1.1}$$

The domain of definition, D, of ψ consists of all the points in a three-dimensional space which lie outside a closed bounded surface Σ. It is assumed that Σ is regular and ψ is regular*.

It is well known that the solution for the same problem associated with the reduced wave equation $(\nabla^2 + k^2)\,\psi = 0$, which obeys SOMMERFELD's "radiation" condition

$$\lim_{R\to\infty} \int\limits_{\Sigma_R} \left| \frac{\partial \psi}{\partial r} - i\,k\,\psi \right|^2 dA = 0 \tag{1.2}$$

where Σ_R is the surface of a sphere of radius R, dA is the element of area of Σ_R and r is the polar coordinate, is uniquely determined by the Dirichlet or Neumann data on the surface Σ. This was rigorously proved by FREUDENTHAL [*2*] (for the two-dimensional case) and others; then by RELLICH [*4*] for any number of dimensions. Later MIRANKER [*5*] considered the case of the Schroedinger wave equation with a locally integrable potential which dies off at infinity as $r^{-\mu}$, $\mu > 3$, and proved that zero Dirichlet (or Neumann) data on the surface Σ imply that a bounded solution vanishes identically outside some sphere. This was sufficient to prove uniqueness when the potential is bounded.

Our main results are:

(i) under suitable restrictions (less strict than MIRANKER's) on V (see Section 2), there does exist a uniqueness theorem for the problem at hand.

(ii) We also prove, incidentally, that the radiation condition implies the boundedness condition: $\psi(\boldsymbol{r}) = O(r^{-1})$, for the solutions considered. Our method of proving this depends on an iteration procedure which resembles the one used by ZEMACH & KLEIN [*6*].

* Regularity of surfaces or functions signifies that the application of GREEN's theorem is permissible. A sufficient condition on ψ is that it be absolutely continuous [*1*]. The characterization of a regular surface is given by KELLOGG [*3*].

Similar results, under different conditions on V, have been obtained independently by KATO [7] by the use of more complicated methods.

Finally, we remark that these considerations fail when $k=0$, and we give a counterexample in this case.

2. Properties of the Potential

We adopt the notation

$$I(\boldsymbol{r}) = \int\limits_D \frac{|V(\boldsymbol{s})|}{|\boldsymbol{r}-\boldsymbol{s}|}\, d\boldsymbol{s},$$

$$I(\boldsymbol{r}, \tau) = \int\limits_\tau \frac{|V(\boldsymbol{s})|}{|\boldsymbol{r}-\boldsymbol{s}|}\, d\boldsymbol{s},$$

where τ is a subregion of D and define

$$\|V\| = (2\pi)^{-1} \operatorname*{Max}_{\boldsymbol{r}\in D} I(\boldsymbol{r}), \tag{2.1}$$

$$\|V\|_\tau = (2\pi)^{-1} \operatorname*{Max}_{\boldsymbol{r}\in D} I(\boldsymbol{r}, \tau). \tag{2.2}$$

The considerations of this paper apply to all potentials $V(\boldsymbol{s})$ which obey the restrictions:

(*i*) $\|V\| < \infty$, (2.3 a)

(*ii*) $I(\boldsymbol{r}) \to 0$ as $r \to \infty$, (2.3 b)

(*iii*) $I(\boldsymbol{r})$ is continuous in $\boldsymbol{r}$, (2.3 c)

(*iv*) $V(\boldsymbol{r})$ is bounded at infinity, (2.3 d)

i.e., there exists constants M and S such that $|V(\boldsymbol{r})| \leqq M$ for $r \geqq S$.

If $V(\boldsymbol{r})$ is bounded everywhere, then (*iii*) is satisfied trivially. These hypotheses embrace potentials of order $O(r^{-2-\varepsilon})$ at infinity and include others of less regular asymptotic behavior. It is clear that $V(\boldsymbol{r})$ must be locally integrable. Condition (*iv*) is used only in Lemma 5 of the next section to establish boundedness of solutions to (1.1). It is unnecessary if boundedness is assumed in advance.

We take note of a few inequalities suggested by the conditions above. The triangle inequality implies that

$$\frac{|\boldsymbol{r}-\boldsymbol{s}|}{|\boldsymbol{r}-\boldsymbol{t}|\,|\boldsymbol{t}-\boldsymbol{s}|} \leqq \frac{1}{|\boldsymbol{r}-\boldsymbol{t}|} + \frac{1}{|\boldsymbol{t}-\boldsymbol{s}|}. \tag{2.4}$$

If (2.4) is multiplied by $|V(t)|$ and integrated over any region τ, we infer, for any $\boldsymbol{r}$, $\boldsymbol{s}$, that

$$\frac{1}{4\pi}\int\limits_\tau \frac{|V(\boldsymbol{t})|}{|\boldsymbol{r}-\boldsymbol{t}|\,|\boldsymbol{t}-\boldsymbol{s}|}\, d\boldsymbol{t} \leqq \frac{\frac{1}{2}(I(\boldsymbol{r},\tau)+I(s,\tau))}{|\boldsymbol{r}-\boldsymbol{s}|} \leqq \frac{\|V\|_\tau}{|\boldsymbol{r}-\boldsymbol{s}|}. \tag{2.5}$$

Alternatively, if σ is a region outside the S sphere defined in (2.3 d), then

$$\begin{aligned}\frac{1}{4\pi}\int\limits_\sigma \frac{|V(\boldsymbol{t})|}{|\boldsymbol{r}-\boldsymbol{t}|\,|\boldsymbol{t}-\boldsymbol{s}|}\, d\boldsymbol{t} &\leqq \int\limits_{\substack{\sigma\\ |\boldsymbol{t}-\boldsymbol{s}|\geqq 1}} + \int\limits_{\substack{\sigma\\ |\boldsymbol{t}-\boldsymbol{s}|\leqq 1}}\\ &\leqq \frac{1}{4\pi}\int\limits_\sigma \frac{|V(\boldsymbol{t})|}{|\boldsymbol{r}-\boldsymbol{t}|}\, d\boldsymbol{t} + \frac{M}{4\pi}\Bigg\{\int\limits_{|\boldsymbol{t}-\boldsymbol{s}|\leqq 1} \frac{d\boldsymbol{t}}{|\boldsymbol{r}-\boldsymbol{t}|^2}\Bigg\}^{\frac{1}{2}} \Bigg\{\int\limits_{|\boldsymbol{t}-\boldsymbol{s}|\leqq 1} \frac{d\boldsymbol{t}}{|\boldsymbol{s}-\boldsymbol{t}|^2}\Bigg\}^{\frac{1}{2}}.\end{aligned}$$

Now, noticing that $\int\limits_{|t-s|\leq 1} \frac{dt}{|r-t|^2}$ has its maximum when $s=r$, we get

$$\frac{1}{4\pi}\int\limits_{\sigma} \frac{|V(t)|}{|r-t|\,|t-s|}\,dt \leq \tfrac{1}{2}\|V\|_\sigma + M. \tag{2.6}$$

Moreover

$$\begin{aligned}\frac{1}{4\pi}\int\limits_{\sigma} \frac{|V(t)|\,dt}{|r-t|\,|s-t|^2} &\leq \frac{1}{4\pi|r-s|}\left\{\int\limits_{\sigma}\frac{|V(t)|}{|t-s|^2}\,dt + \int\limits_{\sigma}\frac{V(t)}{|r-t|\,|s-t|}\,dt\right\}\\ &\leq \frac{(\|V\|_\sigma+2M)}{|r-s|}.\end{aligned} \tag{2.7}$$

Additional properties of the potential are deduced in the lemmas below. In the following lemma, $\sigma(\alpha, r)$ denotes a spherical region of radius α, centered about r and $I(r,\alpha)$, $I(r',\alpha)$ are used in abbreviation of $I(r,\sigma(\alpha,r))$, $I(r',\sigma(\alpha,r'))$ respectively.

Lemma 1. *If α is sufficiently small, then for all $r\in D$, $I(r,\alpha)<1$.*

Proof. Let $\bar\sigma$, $\bar\sigma'$, $\varkappa$ be the complements in D of $\sigma(\alpha,r)$, $\sigma(\alpha,r')$ and $\sigma\cup\sigma'$ respectively, and let τ be the region of points contained in either $\sigma(\alpha,r)$ or $\sigma(\alpha,r')$ but not in both. Suppose $|r-r'|<\frac{1}{2}\alpha$. As $r\to r'$ the volume of $\tau\to 0$, so that

$$\begin{aligned}|I(r,\alpha)-I(r',\alpha)| &\leq |I(r)-I(r')| + |I(r,\bar\sigma)-I(r',\bar\sigma')|\\ &\leq |I(r)-I(r')| + \frac{|r-r'|}{\alpha}\int\limits_{\varkappa}\frac{|V(s)|}{|r-s|}\,ds + \frac{2}{\alpha}\int\limits_{\tau}|V(s)|\,ds\\ &\to 0 \quad \text{as} \quad r\to r'.\end{aligned}$$

Hence, $I(r,\alpha)$ is continuous in r for fixed α. Clearly, for each fixed r, $I(r,\alpha)$ is continuous in α and $\to 0$ as $\alpha\to 0$. Now choose A so large that $I(r)<1$ if $r>A$. It follows that $I(r,\alpha)<1$ for any α if $r>A$. But $I(r,\alpha)\to 0$ uniformly with respect to r as $\alpha\to 0$ if $r\leq A$; hence the lemma is proved.

Let $\{R_i\}$ $1\leq i\leq n$ be a finite sequence of positive numbers, decreasing in magnitude; then D may be divided into a finite number of regions $D_0, D_1, \ldots, D_n$ such that

$$\begin{aligned}&r\in D_0 \quad \text{if} \quad r\in D \quad \text{and} \quad r\geq R_1,\\ &r\in D_i \quad \text{if} \quad r\in D \quad \text{and} \quad R_i>r\geq R_{i+1}, \qquad 1\leq i\leq n-1,\\ &r\in D_n \quad \text{if} \quad r\in D \quad \text{and} \quad R_n>r.\end{aligned}$$

Lemma 2. *The sequences $\{R_i\}$ and $\{D_i\}$ described above may be chosen so that $\|V\|_{D_i}<\frac{1}{2}$, and moreover $R_1>S$.*

Proof. Choose A and α as in Lemma 1, and choose M' so large that

$$\int\limits_{s\geq M'}\frac{|V(s)|}{s}\,ds\leq 1.$$

Then set $R_1=\max\{S, 2A, M'\}$. With R_1 determined, the remaining numbers $R_2, R_3, \ldots$ are to be chosen successively in such a way that

$$\int\limits_{D_i}|V(s)|\,ds\leq\alpha.$$

It is seen that a finite set of R_i suffices. Clearly, if $r > A$, $I(\boldsymbol{r}, D_0) \leqq I(\boldsymbol{r}) \leqq 1$, while if $r \leqq A$,

$$I(\boldsymbol{r}, D_0) \leqq 2 \int\limits_{s \geqq R_1} \frac{|V(\boldsymbol{s})|}{s} d\boldsymbol{s} \leqq 2 \int\limits_{s \geqq M'} \frac{|V(\boldsymbol{s})|}{s} d\boldsymbol{s} \leqq 2.$$

Furthermore, if $i \geqq 1$,

$$I(\boldsymbol{r}, D_i) \leqq \int\limits_{\substack{D_i \\ |\boldsymbol{r}-\boldsymbol{s}| \leqq \alpha}} \frac{|V(\boldsymbol{s})|}{|\boldsymbol{r}-\boldsymbol{s}|} d\boldsymbol{s} + \int\limits_{\substack{D_i \\ |\boldsymbol{r}-\boldsymbol{s}| > \alpha}} \frac{|V(\boldsymbol{s})|}{|\boldsymbol{r}-\boldsymbol{s}|} d\boldsymbol{s} \leqq 1 + \frac{1}{\alpha} \int\limits_{D_i} |V(\boldsymbol{s})| d\boldsymbol{s} \leqq 2.$$

Thus, for all i,

$$\|V\|_{D_i} = (2\pi)^{-1} \operatorname*{Max}_{\boldsymbol{r} \in D} I(\boldsymbol{r}, D_i) < \tfrac{1}{2}.$$

3. Asymptotic Behavior of Solutions to the Wave Equation

Definition. A wave function $\psi(\boldsymbol{r})$ is a regular solution of (1.1) which satisfies the radiation condition (1.2).

The next three lemmas prepare the way for Theorem 1, which demonstrates that a wave function is of order $O(r^{-1})$ at infinity.

Lemma 3. *If $\psi(\boldsymbol{r})$ is a wave function and Σ_s is a sphere about the origin of radius s, then as $s \to \infty$ the following estimates are valid:*

(a)
$$\int\limits_{\Sigma_s} \left| \frac{\partial \psi(\boldsymbol{s})}{\partial s} - i k \psi(\boldsymbol{s}) \right| dA = o(s), \tag{3.1}$$

(b)
$$\int\limits_{\Sigma_s} |\psi(\boldsymbol{s})| \, dA = O(s) \tag{3.2}$$

(c)
$$\int\limits_{\Sigma_s} \left| \frac{\partial \psi(\boldsymbol{s})}{\partial s'} \right| dA = O(s). \tag{3.3}$$

Proof. Equation (3.1) follows directly from (1.2) by use of the Schwarz inequality:

$$\int\limits_{\Sigma_s} \left| \frac{\partial \psi}{\partial s} - i k \psi \right| dA \leqq \left[\int\limits_{\Sigma_s} dA \right]^{\frac{1}{2}} \left[\int\limits_{\Sigma_s} \left| \frac{\partial \psi}{\partial s} - i k \psi \right|^2 dA \right]^{\frac{1}{2}} \leqq (4\pi s)\, o(1) = o(s).$$

To prove (3.2), multiply (1.1) by $\psi^*(\boldsymbol{r})$, take the imaginary part of the result and integrate over the part of D inside the s sphere; then Green's theorem yields

$$\int\limits_{\Sigma_s} \left[\psi^* \frac{\partial \psi}{\partial s} - \psi \frac{\partial \psi^*}{\partial s} \right] dA = \int\limits_{\Sigma} \left[\psi^* \frac{\partial \psi}{\partial n} - \psi \frac{\partial \psi^*}{\partial n} \right] dA$$

where n is the normal into D. Expanding condition (1.2) we get

$$\lim_{s \to \infty} \int\limits_{\Sigma_s} \left\{ \left| \frac{\partial \psi}{\partial s} \right|^2 + k^2 |\psi|^2 + i k \left(\psi^* \frac{\partial \psi}{\partial s} - \psi \frac{\partial \psi^*}{\partial s} \right) \right\} dA = 0.$$

Substituting for the last two terms, we get

$$\lim_{s \to \infty} \int\limits_{\Sigma_s} \left\{ \left| \frac{\partial \psi}{\partial s} \right|^2 + k^2 |\psi|^2 \right\} dA = i k \int\limits_{\Sigma} \left[\psi \frac{\partial \psi^*}{\partial n} - \psi^* \frac{\partial \psi}{\partial n} \right] dA,$$

which implies, since $k \neq 0$, that

$$\int_{\Sigma_s} |\psi|^2 dA \quad \text{and} \quad \int_{\Sigma_s} \left|\frac{\partial \psi}{\partial s}\right|^2 dA$$

remain bounded. Hence

$$\int_{\Sigma_s} |\psi|\, dA \leqq \left[\int_{\Sigma_s} dA\right]^{\frac{1}{2}} \left[\int_{\Sigma_s} |\psi|^2 dA\right]^{\frac{1}{2}} = O(s)\,.$$

Similarly

$$\int_{\Sigma_s} \left|\frac{\partial \psi}{\partial s}\right| dA = O(s)\,.$$

In the work which follows, we make use of the GREEN's function for free space,

$$G(\mathbf{r}, \mathbf{s}) = -\frac{e^{ik|\mathbf{r}-\mathbf{s}|}}{4\pi\,|\mathbf{r}-\mathbf{s}|}\,,$$

which satisfies the equation

$$(\nabla^2 + k^2)\, G(\mathbf{r}, \mathbf{s}) = \delta(\mathbf{r} - \mathbf{s})\,. \tag{3.4}$$

Lemma 4. *Suppose $R > R_1$, where R_1 is defined in Lemma 2, and let D_R be the region enclosed by Σ_R and Σ_{R_1}. Then for $\mathbf{r} \in D_R$*

$$\psi(\mathbf{r}) = \varphi_0(\mathbf{r}) + \int_{D_R} G(\mathbf{r}, \mathbf{s})\, V(\mathbf{s})\, \psi(\mathbf{s})\, d\mathbf{s}\,, \tag{3.5}$$

where $\varphi_0(\mathbf{r})$ is composed of surface integrals over Σ_R and Σ_{R_1}. Moreover

$$\lim_{R\to\infty} \varphi_0(\mathbf{r}) = \int_{\Sigma_{R_1}} \left[G(\mathbf{r}, \mathbf{s}) \frac{\partial}{\partial s} \psi(\mathbf{s}) - \psi(\mathbf{s}) \frac{\partial}{\partial s} G(\mathbf{r}, \mathbf{s})\right] dA_s\,. \tag{3.6}$$

Furthermore, functions $C_1(\mathbf{s})$ and $C_2(\mathbf{s})$ can be chosen such that

$$|\varphi_0(\mathbf{r})| \leqq \int_{\Sigma_R} \left\{\frac{C_1(\mathbf{s})}{|\mathbf{r}-\mathbf{s}|} + \frac{C_2(\mathbf{s})}{|\mathbf{r}-\mathbf{s}|^2}\right\} dA_s \tag{3.7}$$

and

$$\int_{\Sigma_R} C_1(\mathbf{s})\, dA_s = O(R)\,, \qquad \int_{\Sigma_R} C_2(\mathbf{s})\, dA_s = O(R)\,. \tag{3.8}$$

Proof. Multiply (3.4) by $\psi^*(r)$ and (1.1) by $G(\mathbf{r}, \mathbf{s})$, subtract one from the other, integrate over D_R and apply GREEN's theorem. Then (3.5) emerges with

$$\varphi_0(\mathbf{r}) = \int_{\Sigma_R} W(\mathbf{r}, \mathbf{s})\, dA_s - \int_{\Sigma_{R_1}} W(\mathbf{r}, \mathbf{s})\, dA_s\,, \tag{3.9}$$

where

$$W(\mathbf{r}, \mathbf{s}) = \psi(\mathbf{s}) \frac{\partial}{\partial s} G(\mathbf{r}, \mathbf{s}) - G(\mathbf{r}, \mathbf{s}) \frac{\partial}{\partial s} \psi(\mathbf{s})\,.$$

The second integral in (3.9) is seen to be finite for $\mathbf{r}$ within D_R or on Σ_{R_1} (it is the sum of single layer and double layer potentials) and goes to zero as $r \to \infty$; hence it is bounded by some constant C, independent of R, for all $\mathbf{r}$ such that $r \geqq R_1$. To estimate the first integral, we expand $W(\mathbf{r}, \mathbf{s})$:

$$W(\mathbf{r}, \mathbf{s}) = G(\mathbf{r}, \mathbf{s}) \left\{ik \frac{\mathbf{s} - (\mathbf{r}\cdot\mathbf{s})/s}{|\mathbf{r}-\mathbf{s}|} \psi(\mathbf{s}) - \frac{\partial \psi}{\partial s} - \frac{\mathbf{s} - (\mathbf{r}\cdot\mathbf{s})/s}{|\mathbf{r}-\mathbf{s}|^2} \psi(\mathbf{s})\right\}. \tag{3.10}$$

A simple rearrangement of (3.10) shows that, for fixed $\boldsymbol{r}$,

$$|W(\boldsymbol{r},\mathbf{s})| \leqq \left|\frac{\partial\psi}{\partial s} - i\,k\,\psi\right| O\left(\frac{1}{s}\right) + |\psi(\mathbf{s})|\, O\left(\frac{1}{s^2}\right). \tag{3.11}$$

Consequently, by Lemma 3,

$$\begin{aligned}\lim_{R\to\infty} \varphi_0(\boldsymbol{r}) &= -\int_{\Sigma_{R_1}} W(\boldsymbol{r},\mathbf{s})\, dA + \lim_{R\to\infty} \int_{\Sigma_R} W(\boldsymbol{r},\mathbf{s})\, dA \\ &= -\int_{\Sigma_{R_1}} W(\boldsymbol{r},\mathbf{s})\, dA + \lim_{R\to\infty} \left[o(1) + O\left(\frac{1}{R}\right)\right],\end{aligned}$$

and equation (3.6) is proved.

Since

$$|\mathbf{s} - (\boldsymbol{r}\cdot\mathbf{s})/s| \leqq |\mathbf{s} - \boldsymbol{r}|,$$

we have, from (3.10),

$$|W(\boldsymbol{r},\mathbf{s})| \leqq \frac{1}{4\pi\,|\boldsymbol{r}-\mathbf{s}|} \left\{\left|\frac{\partial\psi(\mathbf{s})}{\partial s}\right| + |k\,\psi(\mathbf{s})| + \frac{|\psi(\mathbf{s})|}{|\boldsymbol{r}-\mathbf{s}|}\right\} \tag{3.10a}$$

and

$$|\varphi_0(\boldsymbol{r})| \leqq \int_{\Sigma_R} |W(\boldsymbol{r},\mathbf{s})|\, dA + C \leqq \int_{\Sigma_R} \left\{|W(\boldsymbol{r},\mathbf{s})| + \frac{2C}{4\pi s\,|\boldsymbol{r}-\mathbf{s}|}\right\} dA. \tag{3.12}$$

Equation (3.12) is now seen, with the help of (3.10), to be equivalent to (3.7), and the conditions (3.8) are guaranted by Lemma 3.

Lemma 5. *The wave function $\psi(\boldsymbol{r})$ is bounded in D.*

Proof. The regularity of $\psi(\boldsymbol{r})$ implies boundedness in any bounded region. We shall use (3.5) to set a bound on $\psi(\boldsymbol{r})$ in D_R which does not depend on R, so that, letting $R\to\infty$, we establish the boundedness of $\psi(\boldsymbol{r})$ throughout D. Let (3.5) be iterated, producing the series

$$\psi(\boldsymbol{r}) = \sum_{n-0}^{N} \varphi_n(\boldsymbol{r}) + A_N(\boldsymbol{r}), \tag{3.13}$$

where φ_0 is given by (3.9) and

$$\varphi_n(\boldsymbol{r}) = \int_{D_R} G(\boldsymbol{r},\mathbf{s})\, V(\mathbf{s})\, \varphi_{n-1}(\mathbf{s})\, d\mathbf{s}. \tag{3.14}$$

The remainder term $A_N(\boldsymbol{r})$ is also defined recursively through

$$\begin{aligned} A_0(\boldsymbol{r}) &= \int_{D_R} G(\boldsymbol{r},\mathbf{s})\, V(\mathbf{s})\, \psi(\mathbf{s})\, d\mathbf{s}, \\ A_n(\boldsymbol{r}) &= \int_{D_R} G(\boldsymbol{r},\mathbf{s})\, V(s)\, A_{n-1}(\mathbf{s})\, d\mathbf{s}. \end{aligned}$$

We define

$$\begin{aligned} B_R &= \operatorname*{Max}_{\boldsymbol{r}\in D_R} |\psi(\boldsymbol{r})|, \\ \bar{A}_n &= \operatorname*{Max}_{\boldsymbol{r}\in D_R} |A_n(\boldsymbol{r})|; \end{aligned}$$

then

$$A_n(\boldsymbol{r}) \leqq \frac{\bar{A}_{n-1}}{4\pi} \int_{D_R} \frac{V(\mathbf{s})}{|\boldsymbol{r}-\mathbf{s}|}\, d\mathbf{s} \leqq \left(\frac{1}{2}\, \|V\|_{D_R}\right) \bar{A}_{n-1}.$$

Hence

$$\begin{aligned}\bar{A}_N &\leqq \left(\tfrac{1}{2}\|V\|_{D_R}\right)\bar{A}_{N-1}\\ &\leqq \left(\tfrac{1}{2}\|V\|_{D_R}\right)^N \bar{A}_0\\ &\leqq B_R\left(\tfrac{1}{2}\|V\|_{D_R}\right)^{N+1} \leqq B_R\left(\tfrac{1}{2}\right)^{N+1}\end{aligned}$$

because, by Lemma 2 and the definition of D_R,

$$\|V\|_{D_R} \leqq \|V\|_{D_0} < 1. \tag{3.15}$$

Therefore, we let $N\to\infty$ in (3.13) and obtain

$$\psi(\boldsymbol{r}) = \sum_{n=0}^{\infty}\varphi_n(\boldsymbol{r}).$$

Since R_1 was chosen larger than S defined in (2.3d), inequalities (2.5, 6, 7) are applicable to the region D_R. If we employ them in conjunction with (3.14) and (3.7), we obtain

$$\begin{aligned}|\varphi_1(\boldsymbol{r})| &\leqq \frac{1}{4\pi}\int_{D_R}\frac{1}{|\boldsymbol{r}-\boldsymbol{t}|}|V(\boldsymbol{t})|\,|\varphi_0(\boldsymbol{t})|\,d\boldsymbol{t}\\ &\leqq \int_{\Sigma_R}\frac{K(\boldsymbol{s})}{|\boldsymbol{r}-\boldsymbol{s}|}\,dA_s\end{aligned}$$

where

$$K(\boldsymbol{s}) = C_1(\boldsymbol{s})\,\|V\|_{D_R} + C_2(\boldsymbol{s})\left\{\|V\|_{D_R} + 2M\right\}. \tag{3.16}$$

Thus, by (3.8), $K(s)$ also satisfies

$$\int_{\Sigma_R} K(\boldsymbol{s})\,dA_s = O(R).$$

Invoking (3.14) again, together with (2.5), we have

$$\begin{aligned}|\varphi_2(\boldsymbol{r})| &\leqq \frac{1}{4\pi}\int_{D_R}\frac{1}{|\boldsymbol{r}-\boldsymbol{t}|}|V(\boldsymbol{t})|\,|\varphi_1(\boldsymbol{t})|\,d\boldsymbol{t}\\ &\leqq \|V\|_{D_R}\int_{\Sigma_R}\frac{K(\boldsymbol{s})}{|\boldsymbol{r}-\boldsymbol{s}|}\,dA_s,\end{aligned}$$

and in general, for $n\geqq 1$, the same procedure yields

$$|\varphi_n(\boldsymbol{r})| \leqq \left(\|V\|_{D_R}\right)^{n-1}\int_{\Sigma_R}\frac{K(\boldsymbol{s})}{|\boldsymbol{r}-\boldsymbol{s}|}\,dA.$$

Then

$$\begin{aligned}|\psi(\boldsymbol{r})| &\leqq \sum_{n=0}^{\infty}|\varphi_n(\boldsymbol{r})|\\ &\leqq |\varphi_0(\boldsymbol{r})| + \sum_{n=1}^{\infty}\left(\|V\|_{D_R}\right)^{n-1}\int_{\Sigma_R}\frac{K(\boldsymbol{s})}{|\boldsymbol{r}-\boldsymbol{s}|}\,dA \qquad (3.17)\\ &\leqq |\varphi_0(\boldsymbol{r})| + \left(1-\|V\|_{D_R}\right)^{-1}\int_{\Sigma_R}\frac{K(\boldsymbol{s})}{|\boldsymbol{r}-\boldsymbol{s}|}\,dA.\end{aligned}$$

By (3.7), (3.8), (3.15) and (3.16) we may deduce that, as $R \to \infty$ with r fixed, the right side of (3.17) remains finite. Thus the boundedness of $\psi(\boldsymbol{r})$ in D_0 and hence in D is established.

We are now ready for the principal result of this section.

Theorem 1. *The wave function $\psi(\boldsymbol{r})$ is of order $O(r^{-1})$ at infinity.*

Proof. Let $R \to \infty$ in (3.5), yielding

$$\psi(\boldsymbol{r}) = \psi_0(\boldsymbol{r}) + \int_{D_0} G(\boldsymbol{r}, \mathbf{s})\, V(\mathbf{s})\, \psi(\mathbf{s})\, d\mathbf{s} \tag{3.18}$$

with $\psi_0(\boldsymbol{r}) = \lim_{R\to\infty} \varphi_0(\boldsymbol{r})$ given by (3.6). The integral in (3.18) is absolutely convergent in virtue of (2.3a) and the boundedness of $\psi(\mathbf{s})$. If this equation is iterated, we obtain, for any positive integer N,

$$\psi(\boldsymbol{r}) = \sum_{n=0}^{N} \psi_n(\boldsymbol{r}) + B_N(\boldsymbol{r})$$

where

$$\psi_n(\boldsymbol{r}) = \int_{D_0} G(\boldsymbol{r}, \mathbf{s})\, V(\mathbf{s})\, \psi_{n-1}(\mathbf{s})\, d\mathbf{s},$$

$$B_n(\boldsymbol{r}) = \int_{D_0} G(\boldsymbol{r}, \mathbf{s})\, V(\mathbf{s})\, B_{n-1}(\mathbf{s})\, d\mathbf{s}, \qquad B_0 = \int_{D_0} G(\boldsymbol{r}, \mathbf{s})\, V(\mathbf{s})\, \psi(\mathbf{s})\, d\mathbf{s}.$$

The procedure of Lemma 5, which proved $\lim_{N\to\infty} A_N = 0$ shows, likewise, that

$$\lim_{N\to\infty} \operatorname*{Max}_{\boldsymbol{r}\in D_0} B_N(\boldsymbol{r}) = 0$$

because ψ is bounded in D_0. Therefore, for $\boldsymbol{r} \in D_0$

$$\psi(\boldsymbol{r}) = \sum_{n=0}^{\infty} \psi_n(\boldsymbol{r}).$$

Define

$$\|\psi_n\| = \operatorname*{Max}_{\boldsymbol{r}\in D_0} |r\, \psi_n(\boldsymbol{r})|.$$

Inspection of (3.6) shows that $\|\psi_0\| < \infty$. Furthermore, by (2.5),

$$\begin{aligned}
\|\psi_n\| &\leq \operatorname*{Max}_{\boldsymbol{r}\in D_0} \frac{1}{4\pi} \int_{D_0} \frac{r}{|\boldsymbol{r}-\mathbf{s}|\, s} \|\psi_{n-1}\|\, |V(\mathbf{s})|\, d\mathbf{s}, \\
&\leqq \|V\|_{D_0} \|\psi_{n-1}\| \\
&\leqq (\|V\|_{D_0})^n \|\psi_0\|.
\end{aligned}$$

Therefore, for $\boldsymbol{r} \in D_0$, *i.e.*, for $r \geqq R_1$,

$$\begin{aligned}
r\,|\psi(\boldsymbol{r})| &\leqq \sum_{n=0}^{\infty} (\|V\|_{D_0})^n \|\psi_0\| \\
&\leqq (1 - \|V\|_{D_0})^{-1} \|\psi_0\| < \infty,
\end{aligned}$$

and the proof is concluded.

4. Uniqueness Theorem

Lemma 6. *Let ϑ denote the angle between $\boldsymbol{r}$, $\mathbf{s}$, and suppose $r > s$; then*

$$\left|\frac{e^{ik|\boldsymbol{r}-\boldsymbol{s}|}}{|\boldsymbol{r}-\boldsymbol{s}|} - \frac{e^{ik(r-s\cos\vartheta)}}{r}\right| \leqq \frac{s\left[1+\frac{1}{2}k\,s\right]}{r(r-s)}.$$

Proof. Set $\varphi_1 = |\boldsymbol{r}-\mathbf{s}|$, $\varphi_2 = r - s\cos\vartheta$.

We have, since $(1+a^2)^{\frac{1}{2}} \leqq 1 + \frac{1}{2}a^2$,

$$\begin{aligned}
\varphi_1 &= \sqrt{(r - s\cos\vartheta)^2 + s^2\sin^2\vartheta} \\
&\leqq (r - s\cos\vartheta)\left\{1 + \frac{s^2\sin^2\vartheta}{2(r-s\cos\vartheta)^2}\right\} \\
&\leqq r - s\cos\vartheta + \frac{s^2}{2(r-s)}.
\end{aligned}$$

Therefore, since $\varphi_2 < \varphi_1$,

$$|\varphi_2 - \varphi_1| \leqq \frac{s^2}{2(r-s)}.$$

Also,

$$|\varphi_1 - r| \leqq s.$$

Note that

$$\begin{aligned}
\left|e^{ik\varphi_1} - e^{ik\varphi_2}\right| &= \left|e^{\frac{1}{2}ik(\varphi_1+\varphi_2)}\left\{e^{-ik\frac{(\varphi_1-\varphi_2)}{2}} - e^{ik\frac{(\varphi_1-\varphi_2)}{2}}\right\}\right| \\
&\leqq 2\left|\sin\tfrac{1}{2}k(\varphi_1-\varphi_2)\right| \leqq k|\varphi_1-\varphi_2|.
\end{aligned}$$

Hence

$$\begin{aligned}
\left|\frac{e^{ik\varphi_1}}{\varphi_1} - \frac{e^{ik\varphi_2}}{r}\right| &= \left|e^{ik\varphi_1}\left(\frac{1}{\varphi_1} - \frac{1}{r}\right) + \frac{1}{r}\left(e^{ik\varphi_1} - e^{ik\varphi_2}\right)\right| \\
&\leqq \frac{|\varphi_1 - r|}{\varphi_1 r} + \frac{1}{r}k|\varphi_1-\varphi_2| \\
&\leqq \frac{s\left[1+\frac{1}{2}k\,s\right]}{r(r-s)}
\end{aligned}$$

as required.

Lemma 7. *If*

$$\psi_1(\boldsymbol{r}) = \frac{1}{4\pi}\int\limits_{D_0}\left\{\frac{e^{ik|\boldsymbol{r}-\mathbf{s}|}}{|\boldsymbol{r}-\mathbf{s}|} - \frac{e^{ik[r-s\cos\vartheta]}}{r}\right\}V(\mathbf{s})\,\psi(\mathbf{s})\,d\mathbf{s} \tag{4.1}$$

where $\psi(\mathbf{s})$ is a wave function, then $\psi_1(\boldsymbol{r}) = o(r^{-1})$.

Proof. Suppose that r is large enough that $r > R_1$. Let $\psi_1(\boldsymbol{r}) = A(\boldsymbol{r}) + B(\boldsymbol{r})$, where $A(\boldsymbol{r})$, $B(\boldsymbol{r})$ represent the integral above over the ranges $R_1 \leqq s \leqq r^{\frac{1}{3}}$ and $r^{\frac{1}{3}} \leqq s < \infty$, respectively. By Theorem 1, we can write $|\psi(\mathbf{s})| \leqq C/s$ for some constant C; then by Lemma 6,

$$\begin{aligned}
|A(\boldsymbol{r})| &\leqq \frac{r^{\frac{1}{3}}(1+\frac{1}{2}k\,r^{\frac{1}{3}})}{4\pi r(r-r^{\frac{1}{3}})}\int\limits_{R_1<s<r^{\frac{1}{3}}}\frac{|V(\mathbf{s})|\,C}{s}\,d\mathbf{s} \\
&\leqq \frac{1+\frac{1}{2}k\,r^{\frac{1}{3}}}{r(r^{\frac{2}{3}}-1)}\left[\tfrac{1}{2}C\|V\|\right].
\end{aligned}$$

Furthermore,

$$|B(r)| \leqq \frac{C}{4\pi} \int\limits_{s \geqq r^{\frac{1}{3}}} \left\{ \frac{1}{|\boldsymbol{r}-\boldsymbol{s}|\, s} + \frac{1}{r\, s} \right\} |V(\boldsymbol{s})|\, d\boldsymbol{s}$$

$$\leqq \frac{C}{4\pi r} \int\limits_{s \geqq r^{\frac{1}{3}}} \left\{ \frac{1}{|\boldsymbol{r}-\boldsymbol{s}|} + \frac{2}{s} \right\} |V(\boldsymbol{s})|\, d\boldsymbol{s}$$

$$\leqq \frac{C}{4\pi r} \left[I(\boldsymbol{r}) + 2 \int\limits_{s \geqq r^{\frac{1}{3}}} \frac{|V(\boldsymbol{s})|}{s}\, d\boldsymbol{s} \right].$$

Thus, both $r A(\boldsymbol{r})$ and $r B(\boldsymbol{r})$, and hence $r \psi_1(\boldsymbol{r})$, approach zero as $r \to \infty$.

Theorem 1 shows that a wave function cannot be too large for large values of the argument. We now derive a complementary result to the effect that the wave function cannot be too small at infinity. Use is made of a theorem of RELLICH concerning solutions $\chi(\boldsymbol{r})$ to the reduced wave equation $(\nabla^2 + k^2)\chi(\boldsymbol{r}) = 0$. RELLICH's theorem states that

$$\lim_{R \to \infty} \int\limits_{\Sigma_R} |\chi|^2\, dA \neq 0$$

unless $\chi(\boldsymbol{r})$ vanishes identically.

Theorem 2. *(Analogue of* RELLICH'S *theorem.) If $\psi(\boldsymbol{r})$ is a wave function for* (1.1), *and* $\lim\limits_{R \to \infty} \int\limits_{\Sigma_R} |\psi|^2\, dA = 0$, *then* $\psi(\boldsymbol{r}) \equiv 0$.

Proof. We begin with (3.18) and write $\psi(\boldsymbol{r}) = \psi_1(\boldsymbol{r}) + \psi_2(\boldsymbol{r})$, where $\psi_1(\boldsymbol{r})$ is given by (4.1) and the remainder $\psi_2(\boldsymbol{r})$ is given by

$$\psi_2(\boldsymbol{r}) = \int\limits_{\Sigma_{R_1}} \left[G(\boldsymbol{r}, \boldsymbol{s}) \frac{\partial}{\partial s} \psi(\boldsymbol{s}) - \psi(\boldsymbol{s}) \frac{\partial}{\partial s} G(\boldsymbol{r}, \boldsymbol{s}) \right] dA - \\ - \frac{1}{4\pi} \frac{e^{ikr}}{r} \int\limits_{D_0} e^{-iks\cos\vartheta}\, V(\boldsymbol{s})\, \psi(\boldsymbol{s})\, d\boldsymbol{s}.$$

Lemma 7 and Theorem 1, together with the hypothesis of the present theorem, imply that

$$\lim_{R \to \infty} \int\limits_{\Sigma_R} |\psi_2|^2\, dA = \lim_{R \to \infty} \int\limits_{\Sigma_R} |\psi - \psi_1|^2\, dA$$

$$\leqq \lim_{R \to \infty} \int\limits_{\Sigma_R} |\psi \psi_1^* + \psi_1^* \psi|\, dA$$

$$\leqq 2 \lim_{R \to \infty} \left\{ \int\limits_{\Sigma_R} |\psi|^2\, dA \right\}^{\frac{1}{2}} \left\{ \int\limits_{\Sigma_R} |\psi_1|^2\, dA \right\}^{\frac{1}{2}}$$

$$= 0.$$

But $\psi_2(\boldsymbol{r})$ satisfies $(\nabla^2 + k^2)\psi_2(\boldsymbol{r}) = 0$ for $r > R_1$ since $\frac{e^{ikr}}{r}$, $G(\boldsymbol{r}, \boldsymbol{s})$ and its derivatives do so (for $r \neq 0$ and $\boldsymbol{s} \in \Sigma_{R_1}$). Hence RELLICH's theorem implies $\psi_2(\boldsymbol{r}) \equiv 0$, so that the revised equation for the wave function for $r \in D_0$ reads

$$\psi(\boldsymbol{r}) = -\frac{1}{4\pi} \int\limits_{D_0} \left\{ \frac{e^{ik|\boldsymbol{r}-\boldsymbol{s}|}}{|\boldsymbol{r}-\boldsymbol{s}|} - \frac{e^{ik(r - s\cos\vartheta)}}{r} \right\} V(\boldsymbol{s})\, \psi(\boldsymbol{s})\, d\boldsymbol{s}.$$

With the notation $\|\psi\| = \operatorname*{Max}_{\boldsymbol{r}\in D_0} r\,|\psi(\boldsymbol{r})| < \infty$, we get

$$\|\psi\| \leqq \frac{1}{4\pi}\|\psi\| \int\limits_{D_0} \left| \frac{r}{s\,|\boldsymbol{r}-\boldsymbol{s}|} + \frac{1}{s} \right| |V(\boldsymbol{s})|\, d\boldsymbol{s}$$

$$\leqq \|\psi\| \cdot \tfrac{3}{2} \|V\|_{D_0} \leqq \tfrac{3}{4} \|\psi\|$$

because $\|V\|_{D_0} < \frac{1}{2}$. Therefore $\|\psi\| = 0$ and $\psi(\boldsymbol{r}) = 0$ for $r \in D_0$. Suppose that the vanishing of $\psi(\boldsymbol{r})$ for $\boldsymbol{r}\in D_i$ $i \leqq m$ has been proven. We shall show that ψ vanishes in D_{m+1}. We suppose that the origin of the coordinate system lies inside Σ, so that $r \neq 0$ in D. Applying GREEN's theorem to the region D_{m+1}, when $\boldsymbol{r}\in D_{m+1}$ we obtain a relation identical to (3.18) except that Σ_{R_0} and D_0 are replaced by $\Sigma_{R_{m+2}}$ and D_{m+1}. We proceed precisely as in the above paragraph, apply RELLICH's theorem to the region $r > R_{m+2}$, and deduce in an analogous manner that $\psi(\boldsymbol{r}) = 0$ when $\boldsymbol{r}\in D_{m+1}$. Therefore $\psi(\boldsymbol{r})$ vanishes in all D_i, and the proof is completed.

Theorem 3. *(Dirichlet-Neumann uniqueness theorem.) Let $\psi'(\boldsymbol{r})$ and $\psi''(\boldsymbol{r})$ be regular in D, satisfy the Schroedinger wave equation* (1.1), *the radiation condition and the boundary conditions*

(a) $$\psi'(\boldsymbol{r}) = \psi''(\boldsymbol{r}) = a(\boldsymbol{r}) \qquad \textit{(Dirichlet condition)}$$

or

(b) $$\frac{\partial\psi'(\boldsymbol{r})}{\partial n} = \frac{\partial\psi''(\boldsymbol{r})}{\partial n} = b(\boldsymbol{r}) \qquad \textit{(Neumann condition)}$$

on the surface Σ of D. Then $\psi'(\boldsymbol{r}) \equiv \psi''(\boldsymbol{r})$.

Proof. Set $\psi(\boldsymbol{r}) = \psi'(\boldsymbol{r}) - \psi''(\boldsymbol{r})$; then $\psi(\boldsymbol{r})$ is a wave function which obeys a zero Dirichlet or Neumann condition on Σ.

Applying GREEN's theorem to $\psi(\boldsymbol{r})$, $\psi^*(\boldsymbol{r})$ in the region between Σ and the spherical surface Σ_R, we have

$$\int\limits_{\Sigma_R} \left[\psi \frac{\partial\psi^*}{\partial n} - \psi^* \frac{\partial\psi}{\partial n}\right] dA = \int\limits_{\Sigma} \left[\psi \frac{\partial\psi^*}{\partial n} - \psi^* \frac{\partial\psi}{\partial n}\right] dA = 0. \tag{4.2}$$

Expanding (1.2), we see that (4.2) leads to

$$\lim_{R\to\infty} \int\limits_{\Sigma_R} |\psi|^2\, dA = \lim_{R\to\infty} \int\limits_{\Sigma_R} \left|\frac{\partial\psi}{\partial n}\right|^2 dA = 0.$$

Therefore, by the previous theorem $\psi(\boldsymbol{r}) \equiv 0$, and $\psi'(\boldsymbol{r}) \equiv \psi''(\boldsymbol{r})$.

Corollary. *A function which satisfies the radiation condition, and which is a regular solution of* (1.1) *in the whole space, is identically zero.*

A counter-example. In the special case $k = 0$, Theorem 2 (and hence Theorem 3) ceases to be true. We give a counter-example to the corollary above for this case:

Let

$$V(\boldsymbol{r}) = \begin{cases} \dfrac{-30}{5-3r^2} & r \leqq 1 \\ 0 & r > 1 \end{cases}$$

$$\psi(\boldsymbol{r}) = \begin{cases} r[5-3r^2]\cos\vartheta & r \leqq 1 \\ \dfrac{2}{r^2}\cos\vartheta & r \geqq 1\,. \end{cases}$$

Then ψ satisfies the radiation condition, is regular in the whole space and $[\nabla^2 - V]\psi = 0$, and yet $\psi \not\equiv 0$.

This work was supported in part by the United States Air Force under contract No. AF 49(638)-327 monitored by the AF-Office of Scientific Research of the Air Research and Development Command and the Office of Naval Research under Contract No. r-222(60) (NR 041—221).

References

[1] Titchmarsh, E. C.: Eigenfunctions Expansions, Part II, pp. 325—326. Oxford 1958.
[2] Freudenthal, H.: Über ein Beugungsproblem aus der elektromagnetischen Lichttheorie. Compositio Math. **6**, 221—227 (1938).
[3] Kellogg, O.D.: Foundations of Potential Theory, p. 105. Berlin: Springer 1929.
[4] Rellich, F.: Über das asymptotische Verhalten der Lösungen von $\Delta u + \lambda u = 0$ in unendlichen Gebieten. Jahresber. d. D.M.V. **53**, 57—65 (1943).
[5] Miranker, W.: The reduced wave equation in a medium with a variable index of refraction. Comm. Pure and Applied Math. **10**, 491—502 (1957).
[6] Zemach, C., & A. Klein: The Born expansion in non-relativistic quantum theory. Nuovo Cimento, Ser. X, **10**, 1078—1089 (1958).
[7] Kato, T.: Growth properties of solutions to the reduced wave equation with a variable coefficient. Comm. Pure and Applied Math. **12**, 403—425 (1959).

University of California
Berkeley, California

(Received December 21, 1959)

Mildly Nonlinear Parabolic Equations with Application to Flow of Gases through Porous Media

AVNER FRIEDMAN

Communicated by J. SERRIN

Introduction

The study of the motion of a homogeneous gas in a porous medium, when the equation of state is

$$\gamma = \gamma_0 p^m \quad (\gamma = \text{density},\ p = \text{pressure},\ \gamma_0 \text{ constant}), \tag{0.1}$$

is reduced (see [7]) to the study of the nonlinear parabolic equation

$$\Delta \gamma^{(1+m)/m} = \frac{1}{k}(1+m) f \mu \gamma_0^{1/m} \frac{\partial \gamma}{\partial t} \tag{0.2}$$

where Δ is the 3-dimensional Laplacian, k = permeability of the medium, f = porosity of the medium and μ = viscosity of the gas. For isothermic flows $m=1$, and for adiabatic flows $0<m<1$. In deriving (0.2), it is assumed, as is most commonly the case, that the external forces may be ignored.

Equation (0.2) is a particular case of the class of nonlinear equations

$$\frac{\partial u}{\partial t} = L[\psi(x,u)] + f(x,t) \tag{0.3}$$

where $\psi(x,0) \geqq 0$ and $\partial\psi/\partial u > 0$ if $u>0$, and where L is a linear elliptic operator with coefficients depending on $x=(x_1,\dots,x_n)$. As will be shown later on, (0.3) is equivalent to the equation

$$\frac{\partial v}{\partial t} = k(x,v)\, L\, v + k(x,v) f(x,t) \tag{0.3'}$$

where $k(x,v)>0$ if $v>0$. General mildly nonlinear equations

$$\frac{\partial u}{\partial t} = \sum a_{ij}(x,t,u) \frac{\partial^2 u}{\partial x_i\, \partial x_j} + \sum b_i(x,t,u) \frac{\partial u}{\partial x_i} + f(x,t,u) \tag{0.4}$$

have been discussed in the literature. The first boundary value problem has been solved in cylindrical domains by OLAINIK & VENTSEL [*11*] for $n=1$ and by VENTSEL [*15*] and LADYZHENSKAYA [*5*] for general n. However, for $n>1$ they assume that the height of the cylinder is appropriately small.

In this paper we are interested in existence for all $t>0$. We first consider equation (0.4) and prove (using Schauder type estimates) that if we have an *a priori* Hölder inequality (or an *a priori* bound on $\partial u/\partial t$), then existence and uniqueness can be established for all $t>0$. Next we consider equation (0.3) and prove, under some restrictions on k, f and the boundary data, that there is an *a priori* bound on $\partial u/\partial t$.

1. Mildly nonlinear parabolic equations

Let B be an n-dimensional bounded domain with boundary ∂B. For any $\varrho>0$, denote by D_ϱ the topological product of B by $\{0<t<\varrho\}$ and by S_ϱ the topological product of ∂B by $\{0\leqq t<\varrho\}$. Set $D=D_\infty$, $S=S_\infty$. By $\overline{D}_\varrho$ we denote the closure of D_ϱ. We recall that a function $v(x,t)$ is said to be Hölder continuous (exponent α) in a set G if

$$H_\alpha^G(v)\equiv \underset{(x,t)\in G,\,(x^0,t^0)\in G}{\mathrm{l.u.b.}} \frac{|v(x,t)-v(x^0,t^0)|}{|x-x^0|^\alpha+|t-t^0|^{\alpha/2}} \quad \text{is finite.}$$

We write

$$|v|_0^G=\underset{G}{\mathrm{l.u.b.}}\,|v(x,t)|,\qquad |v|_\alpha^G=|v|_0^G+H_\alpha^G(v)$$

$$|v|_{2+\alpha}^G=|v|_\alpha^G+\sum\left|\frac{\partial v}{\partial x_i}\right|_\alpha^G+\sum\left|\frac{\partial^2 v}{\partial x_i\,\partial x_j}\right|_\alpha^G+\left|\frac{\partial v}{\partial t}\right|_\alpha^G.$$

We say that v is of class C^k in G if $|v|_k^G<\infty$. We denote by $|v|_k^\varrho$ the k-norm of v in D_ϱ.

We shall consider the mildly nonlinear parabolic equation

$$(1.1)\qquad L u\equiv\sum_{i,j=1}^n a_{ij}(x,t,u)\frac{\partial^2 u}{\partial x_i\,\partial x_j}+\sum_{i=1}^n b_i(x,t,u)\frac{\partial u}{\partial x_i}-\frac{\partial u}{\partial t}=f(x,t,u)$$

in D, with the initial and boundary conditions

$$(1.2)\qquad u(x,0)=h(x)\ \text{ on } B,\qquad u(x,t)=g(x,t)\ \text{ on } S.$$

We shall need the following assumptions:

(A_1) L is of parabolic type; that is, for every real vector $\xi\neq 0$ and for every $(x,t)\in\overline{D}$, $\quad -\infty<u<\infty,\quad \sum a_{ij}(x,t,u)\,\xi_i\xi_j>0$.

(A_2) f and the coefficients of L have two continuous derivatives for $(x,t)\in\overline{D}$, $-\infty<u<\infty$.

(A_3) The boundary ∂B is of class $C^{2+\alpha}$, for some $0<\alpha<1$.

(A_4) There exists a function $\Phi(x,t)$ which coincides with $h(x)$ on B and with $g(x,t)$ on S such that $|\Phi|_{2+\alpha}^\varrho<\infty$ for any $\varrho>0$.

Theorem 1. *Let the assumptions* $(A_1)-(A_4)$ *be satisfied. Assume that for every* $\varrho>0$ *there exists a constant* M_ϱ *depending only on* ϱ, L, f, Φ, B *such that whenever a solution* u *of* (1.1), (1.2) *exists in* D_ϱ, *it necessarily satisfies the inequality*

$$(1.3)\qquad |u|_\alpha^\varrho\leqq M_\varrho.$$

Then there exists a unique solution of (1.1), (1.2) *in* D.

Proof. It is enough to prove existence and uniqueness in D_μ for every fixed μ. The proof is given in three steps: (i) existence in D_ϱ for small ϱ, (ii) uniqueness in D_ϱ for any ϱ, and (iii) existence in D_μ by repeated application of (i), making use of (ii) and the *a priori* inequality (1.3). Since, by assumption, u is *a priori* bounded, we may change the definition of f and the coefficients of L for $|u|$ large without affecting the assertion of the theorem. Hence, without loss of generality, we may assume in what follows that L is uniformly parabolic and that the second derivatives of f and of the coefficients of L are uniformly bounded for $(x,t)\in\bar{D}_\mu$, $-\infty<u<\infty$.

(i) **Existence in D_ϱ, ϱ small.** Denote by Z^ϱ the Banach space of functions $v(x,t)$ defined in D_ϱ and having a finite norm $|v|_\alpha^\varrho$. We define a transformation $w=Tv$ as follows: Replace in f and in the coefficients of L the variable u by $v(x,t)$ and let w by the solution of the modified system (1.1), (1.2). By [3], w exists and is uniquely determined. Furthermore, since the coefficients of the modified equation are Hölder continuous (exponent α), we have

$$(1.4)\qquad |w|_{2+\alpha}^\varrho\leqq K_0(N)\,(|\Phi|_{2+\alpha}^\varrho+|f|_\alpha^\varrho)\equiv K_1(N)\quad\text{provided}\quad |v|_\alpha^\varrho\leqq N$$

where K_i depend only on N, ϱ, L, B and K_1 also on f. We may however take K_1 sufficiently large such that it is independent of ϱ, $0<\varrho\leqq\mu$ (indeed, we can just take the K_1 of (1.4) which corresponds to $\varrho=\mu$).

We proceed to show that for ϱ sufficiently small T has a fixed point $v=Tv$. By SCHAUDER's fixed point theorem [13] it is sufficient to show that T maps a closed convex set Z_N^ϱ of Z defined by

$$(1.5)\qquad |v|_\alpha^\varrho\leqq N$$

into a compact subset of itself, and that T is continuous.

T maps Z_N^ϱ into itself. For any $0<t<t^0<\varrho$ we have

$$(1.6)\qquad \frac{|w(x,t)-w(x,t^0)|}{|t^0-t|^{\alpha/2}}\leqq(t^0-t)^{1-\alpha/2}\left|\frac{\partial w}{\partial t}\right|_0^\varrho\leqq K_1(N)\,\varrho^{1-\alpha/2}$$

where we have made use of (1.4). Also,

$$\left|\frac{\partial}{\partial t}w(x,t)-\frac{\partial}{\partial t}w(x^0,t)\right|\leqq|x-x^0|^\alpha\left|\frac{\partial w}{\partial t}\right|_\alpha^\varrho\leqq K_1(N)\,|x-x^0|^\alpha.$$

By integration we then obtain

$$(1.7)\qquad \frac{|w(x,t)-w(x^0,t)|}{|x-x^0|^\alpha}\leqq\frac{|h(x)-h(x^0)|}{|x-x^0|^\alpha}+K_1(N)\,\varrho.$$

Combining (1.6) and (1.7), we conclude (if $\varrho\leqq1$)

$$(1.8)\qquad H_\alpha^\varrho(w)\leqq H_\alpha(h)+2K_1(N)\,\varrho^{1-\alpha/2}.$$

By the maximum principle [9] we have

$$(1.9)\qquad |w|_0^\varrho\leqq N_0$$

where N_0 is independent of v and can also be taken to be independent of ϱ, $0<\varrho\leqq\mu$. Combining (1.8) and (1.9), we conclude that if we choose

$$N = 2M_\mu + N_0 + 1 \qquad (M_\mu \text{ appears in } (1.3)), \tag{1.10}$$

then T maps Z_N^ϱ into itself provided

$$2K_1(N)\,\varrho^{1-\alpha/2} \leqq 1. \tag{1.11}$$

(Remark: $N = M_\mu + N_0 + 1$ would suffice since $H_\alpha(h) \leqq M_\mu$; the reason for the definition of N by (1.10) is that it will slightly facilitate the calculations in step (iii).)

Clearly, T maps Z_N^ϱ into a compact subset.

T is continuous on Z_N^ϱ. Let v, v^0 belong to Z_N^ϱ. Subtracting the equation satisfied by $w^0 = Tv^0$ from that satisfied by $w = Tv$ and setting $y = w - w^0$, we obtain

$$\sum a_{ij}(x,t,v)\frac{\partial^2 y}{\partial x_i\,\partial x_j} + \sum b_i(x,t,v)\frac{\partial y}{\partial x_i} - \frac{\partial y}{\partial t} = F(x,t) \tag{1.12}$$

where

$$\begin{aligned} F = \sum\,[a_{ij}(x,t,v^0) - a_{ij}(x,t,v)]\,\frac{\partial^2 w^0}{\partial x_i\,\partial x_j} + \sum\,[b_i(x,t,v^0) - {}\\ - b_i(x,t,v)]\,\frac{\partial w^0}{\partial x_i} + [f(x,t,v) - f(x,t,v^0)]. \end{aligned} \tag{1.13}$$

Furthermore, $y=0$ on B and on S_ϱ. Using the estimate (1.4) we get

$$H_\alpha^\varrho(y) \leqq K(N)\,|F|_\alpha^\varrho. \tag{1.14}$$

By the maximum principle,

$$|y|_0^\varrho \leqq \text{const.}\,|F|_0^\varrho \leqq \text{const.}\,|v - v^0|_0^\varrho. \tag{1.15}$$

Using the mean value theorem, after some elementary calculations we obtain

$$H_\alpha^\varrho(F) \leqq \text{const.}\,(|v^0 - v|_\alpha^\varrho)^{\frac{1}{2}}.$$

Substituting this inequality and (1.15) into (1.14), we conclude that T is continuous.

(ii) **Uniqueness in $\boldsymbol{D_\rho}$, $\boldsymbol{\rho\leqq\mu}$.** Let u, u^0 be solutions of (1.1), (1.2) in a domain D_ϱ, $\varrho\leqq\mu$, and assume that u is of class C^α in $\overline{D}_\varrho$ (later we shall construct a C^α solution). It is then easy to prove that $u\equiv u^0$. Indeed, subtracting the differential equations satisfied by u^0 from that satisfied by u, we obtain for $y = u - u^0$ an equation of the form (1.12) with v replaced by $u^0(x,t)$ and with F defined by (1.13) with v, v^0, w^0 replaced by u^0, u, u respectively. It is then immediately seen that $F = Ey$ where E is a continuous function of (x,t) in $\overline{D}_\varrho$. Since $y=0$ on B and on S_ϱ, we conclude, by the maximum principle, that $y = u - u^0 \equiv 0$ in D_ϱ.

(iii) **Existence in D_μ.** In this proof we shall make use of the assumption (1.3). We wish to repeat the construction of (i) step-by-step. We shall show that each step can be carried out in a cylinder with the same height ϱ as used in (i). The uniqueness proved in (ii) assures that a step-by-step construction yields a unique solution in the whole domain D_μ.

We thus suppose that a solution has been constructed in D_σ for some $0<\sigma<\mu$, and we shall extend it to $D_{\sigma+\varrho}$ (assuming, as we may, that $\sigma+\varrho\leqq\mu$).

Let $Z^{\sigma+\varrho}$ be the Banach space of functions $v(x,t)$ which coincide in D_σ with the solution $u(x,t)$ and for which $|v|_\alpha^{\sigma+\varrho}<\infty$. We define $w=Tv$ as follows: Replace in f and in the coefficients of (1.1) the variable u by $v(x,t)$, and let w be the solution of the modified system (1.1), (1.2). By (ii), $w\equiv u$ in D_σ and, hence, w also belongs to $Z^{\sigma+\varrho}$. To show that T has a fixed point it is sufficient to show that T maps a set $Z_N^{\sigma+\varrho}$ (for some $N>0$) defined by $|v|_\alpha^{\sigma+\varrho}\leqq N$ into a compact subset of itself and that T is continuous. The continuity of T is proved as in (i). The compactness of $T(Z_N^{\sigma+\varrho})$ is immediate. It thus remains to show that T maps $Z_N^{\sigma+\varrho}$ into itself.

We have

$$\frac{|w(x,t)-w(x,t^0)|}{|t-t^0|^{\alpha/2}}\leqq K_1(N)\,\varrho^{1-\alpha/2} \quad \text{provided } \sigma<t<t^0<\sigma+\varrho. \tag{1.16}$$

Also, recalling that $w(x,\sigma)=u(x,\sigma)$, we get as in (i),

$$\frac{|w(x,t)-w(x^0,t)|}{|x-x^0|^{\alpha}}\leqq\frac{|u(x,\sigma)-u(x^0,\sigma)|}{|x-x^0|^{\alpha}}+K_1(N)\,\varrho \quad \text{provided } \sigma\leqq t<\sigma+\varrho. \tag{1.17}$$

Using (1.3) to estimate the right side of (1.17) and combining the inequality thus obtained with (1.16), we get

$$H_\alpha^G(w)\leqq M_\mu+2K_1(N)\,\varrho^{1-\alpha/2} \quad \text{where} \quad G=D_{\sigma+\varrho}-D_\sigma. \tag{1.18}$$

Since, by (1.3), $H_\alpha^\sigma(w)\leqq M_\mu$, we conclude

$$H_\alpha^{\sigma+\varrho}(w)\leqq 2M_\mu+2K_1(N)\,\varrho^{1-\alpha/2}. \tag{1.19}$$

Using (1.19), (1.9), we find that if we define N (as in (i)) by (1.10), then T maps $Z_N^{\sigma+\varrho}$ into itself provided ϱ is restricted by (1.11). Thus ϱ can be chosen to be independent of σ, and the proof is completed.

To be able to prove existence for (1.1), (1.2) in D, one needs, via Theorem 1, to establish the *a priori* inequality (1.3). An *a priori* bound on $|u|_0^\varrho$ can be obtained in some cases by the use of the maximum principle. An *a priori* inequality on $H_\alpha^\varrho(u)$ is much deeper. This question would be solved if one could establish Hölder continuity (in the whole domain) of solutions of linear parabolic equations, where the Hölder coefficient and exponent depend *only* on bounds on the coefficients, bound on the solution, and the modulus of parabolicity of the equation. This has been solved, however, only for a half strip when the elliptic part of the equations is of the form $\sum\frac{\partial}{\partial x_i}\left(a_{ij}\frac{\partial}{\partial x_j}\right)$ (J. NASH [8]).

In some cases however, an *a priori* bound can be derived on $\partial u/\partial t$. This will be shown in § 2. Hence arises the usefulness of the following theorem.

Theorem 2. *Let the assumptions* $(A_1)-(A_4)$ *be satisfied. Assume that for any* $\varrho>0$ *there exists a constant* M_ϱ *depending only on* ϱ, L, f, Φ, B *such that whenever a solution* u *of* (1.1), (1.2) *exists in* D_ϱ, *it necessarily satisfies the inequality*

$$|u|_0^\varrho + \left|\frac{\partial u}{\partial t}\right|_0^\varrho \leqq M_\varrho . \tag{1.20}$$

Assume finally that the matrix $(a_{ij}(x, t, u))$ *satisfies the* K_ε*-condition of* CORDES [*1*; p. 292]. *Then there exists a unique solution of* (1.1), (1.2) *in* D.

Remark. The K_ε-condition states that the eigenvalues $\lambda_1, \ldots, \lambda_n$ of (a_{ij}) belong to some circular cone in the domain $\lambda_1>0, \ldots, \lambda_n>0$, having the origin for vertex. It is satisfied in the special case of $a_{ij}(x, t, u) - k(u)\,a_{ij}(x, t)$ provided $k \geqq \text{const.} > 0$, and (a_{ij}) is a continuous positive matrix.

Proof. By [*1*] (using (1.20)) we get an *a priori* bound on the Hölder continuity of u and $\partial u/\partial x_i$ with respect to x. Combining it with (1.20), we obtain an *a priori* bound on $|u|_\alpha^\varrho$.

Remark. In some cases, when the equation is written in the form

$$\sum \frac{\partial}{\partial x_i}\left(a_{ij}\frac{\partial u}{\partial x_j}\right) + \sum b_i \frac{\partial u}{\partial x_i} - \frac{\partial u}{\partial t} = f$$

the K_ε-condition can be removed, and instead of using [*1*] we can use the *a priori* estimate of DE GIORGI [*2*] and NASH [*8*] as generalized by MORREY [*6*].

2. Generalized flow of gases through porous media

In this chapter we shall consider a special case of (1.1), namely,

$$\frac{\partial u}{\partial t} = k(x, u)\,[L_0 u + f(x, t)] \tag{2.1}$$

where

$$L_0 u \equiv \sum a_{ij}(x)\,\frac{\partial^2 u}{\partial x_i\,\partial x_j} + \sum b_i(x)\,\frac{\partial u}{\partial x_i} + c(x)\,u . \tag{2.2}$$

The boundary conditions are as in (2.1). We shall first prove existence, uniqueness and stability, and then we shall show that the equation of flow of gas through a porous medium reduces to a special case of (2.1). For simplicity we first formulate some assumptions.

(A_5) L_0 is uniformly elliptic in $\bar{B}$, its coefficients are Hölder continuous (exponent α) in $\bar{B}$, and $c(x) \leqq 0$.

(A_6) $k(x, u)$ has two continuous derivatives for $x \in \bar{B}$, $0<u<\infty$, and $k(x, u)>0$ $\partial k(x, u)/\partial u \geqq 0$.

(A_7) $f(x, t)$ and $\partial f(x, t)/\partial t$ are Hölder continuous (exponent α) in each D_ϱ, $\varrho>0$, and $f \geqq 0$, $\partial f/\partial t \leqq 0$.

(A_8) $h(x)>0$ and $L_0 h(x) + f(x, 0) \leqq 0$ on B; $g(x, t)>0$ and $\partial g(x, t)/\partial t \leqq 0$ on S.

Theorem 3. *Let the assumptions* $(A_3)-(A_8)$ *be satisfied. Then there exists a unique solution of* (2.1), (1.2). *The solution belongs to* $C^{2+\alpha}$ *in each* D_ϱ, $\varrho>0$.

Proof. The proof consists of two steps: (i) Deriving bounds on u, and (ii) deriving bounds on $\partial u/\partial t$.

(i) **Bounds on u.** Consider the function

$$\varphi(x)=\exp\{\lambda R\}-\exp\{\lambda(x_1-x_1^0)\} \qquad (\lambda>0,\ R>0). \tag{2.3}$$

We take x_1^0 and R such that $0\leqq x_1-x_1^0<R$ for $x=(x_1,\dots,x_n)$ in $\overline{B}$, and then choose λ such that $L_0\varphi(x)<-\gamma$ in B, where γ is some positive constant. We compare, in D_ϱ, the solution u with

$$\psi(x)=\gamma^{-1}|f|_0^\varrho\,\varphi(x). \tag{2.4}$$

We have

$$k(x,u)\,L_0\psi-\frac{\partial\psi}{\partial t}\leqq -k(x,u)\,|f(x,t)| \text{ in } D_\varrho.$$

Hence, the function $v=\psi-u$ satisfies in D_ϱ

$$k(x,u)\,L_0 v-\frac{\partial v}{\partial t}\leqq 0.$$

On the boundary $B+S_\varrho$ $v\geqq -u\geqq -\max(|g|_0^\varrho,|h|_0)\equiv -C_1$. Using the maximum principle [*9*], we conclude that $v\geqq -C_1$ in D_ϱ, or $u\leqq\psi+C_1$.

We next need to show that u is bounded (in D_ϱ) from below by a *positive* constant (since $k(x,u)$ need not be defined for $u=0$). Since $f\geqq 0$,

$$\frac{\partial u}{\partial t}\geqq k(x,u)\,L_0 u \text{ in } D_\varrho. \tag{2.5}$$

Since $u>0$ in D_ϱ, the right side of (2.5) is elliptic. The function $\tilde{u}(x,t)=u(x,t)$ $\exp\{-\nu t\}$ $(\nu>0)$ satisfies

$$\frac{\partial\tilde{u}}{\partial t}\geqq k(x,u)\,L_0\tilde{u}+\nu\tilde{u}.$$

Since we have already proved that $u<\psi+C_1$, it follows that $k(x,u)$ is bounded from above. Hence, if ν is sufficiently large (independently of u), then

$$\frac{\partial\tilde{u}}{\partial t}\geqq k(x,u)\,[L_0-c(x)]\,\tilde{u} \text{ in } D_\varrho.$$

We can now apply the maximum principle and obtain

$$\tilde{u}(x,t)\geqq\exp\{-\nu\varrho\}\,C_2 \quad\text{where}\quad C_2=\min(\operatorname*{g.l.b.}_{B} h,\ \operatorname*{g.l.b.}_{S_\varrho} g),$$

and the proof is completed.

(ii) **Bounds on $\partial u/\partial t$.** Using (i), it follows that we may change the definition of $k(x,u)$, for u large, without any loss in generality of the proof. We thus may assume in what follows that k and its first two derivatives are bounded uniformly for $x\in\overline{B}$, $-\infty<u<\infty$ and that k is uniformly bounded from below by a positive constant. We shall need a lemma essentially due to WESTPHAL [*16*].

Lemma 1. *Let $v(x, t)$, $w(x, t)$ be continuous functions in $\overline{D}_\varrho$, having in D_ϱ second continuous x-derivatives and first continuous t-derivatives. Let $F(x, t, z, z_i, z_{ij})$ $(i, j=1, \ldots, n)$ be a continuous function in the closure of the domain E defined by*

$$(x, t) \in D_\varrho, \quad z \in \big(v(x, t), w(x, t)\big), \quad z_i \in \left(\frac{\partial v(x, t)}{\partial x_i}, \frac{\partial w(x, t)}{\partial x_i}\right),$$
$$z_{ij} \in \left(\frac{\partial^2 v(x, t)}{\partial x_i \, \partial x_j}, \frac{\partial^2 w(x, t)}{\partial x_i \, \partial x_j}\right)$$

where (q, r) denotes the open interval between q and r, and assume that $(\partial F/\partial z_{ij})$ is a continuous positive matrix in E. Assume, finally, that

$$\frac{\partial v}{\partial t} \geqq F\left(x, t, v, \frac{\partial v}{\partial x_i}, \frac{\partial^2 v}{\partial x_i \, \partial x_j}\right) \quad \text{for} \quad (x, t) \in D_\varrho,$$
$$\frac{\partial w}{\partial t} < F\left(x, t, w, \frac{\partial w}{\partial x_i}, \frac{\partial^2 w}{\partial x_i \, \partial x_j}\right) \quad \text{for} \quad (x, t) \in D_\varrho,$$

and that

$$v > w \quad \text{on} \quad \overline{B} + S_\varrho.$$

Then $v(x, t) > w(x, t)$ in D_ϱ.

The lemma holds also for noncylindrical domains.

We now denote by $z(x, t)$ the solution of the systems

$$\text{(2.6)} \qquad \frac{\partial z}{\partial t} = k(x, u) \, L_0 z + \frac{\partial k(x, u)}{\partial u} \frac{1}{(k(x, u))^2} \left(\frac{\partial u}{\partial t}\right)^2 + k(x, u) \frac{\partial f(x, t)}{\partial t}$$

$$\text{(2.7)} \qquad z = \frac{\partial g}{\partial t} \text{ on } S, \quad z = k(x, 0) \, [L_0 h(x) + f(x, 0)] \text{ on } B.$$

The existence of z follows by [3]. Next we consider the differential system satisfied by $u^h(x, t) = [u(x, t+h) - u(x, t)]/h$ and compare it with (2.6), (2.7). Using the maximum principle, we then find that $|u^h - z|_0^\varrho \to 0$ as $h \to 0$. Hence, $z = \partial u/\partial t$. The equation (2.6) then takes the form

$$\text{(2.8)} \qquad \frac{\partial z}{\partial t} = k(x, u) \, L_0 z + \frac{\partial k(x, u)}{\partial u} \frac{1}{(k(x, u))^2} z^2 + k(x, u) \frac{\partial f(x, t)}{\partial t}.$$

Note, by (2.7) and the assumption (A_8), that

$$\text{(2.9)} \qquad z(x, 0) \leqq 0 \text{ on } B, \quad z(x, t) \leqq 0 \text{ on } S_\varrho.$$

We proceed to estimate $z = \partial u/\partial t$ from below. Since $\partial k/\partial u \geqq 0$, we have

$$\text{(2.10)} \qquad \frac{\partial z}{\partial t} \geqq k(x, u) \, L_0 z + k(x, u) \frac{\partial f(x, t)}{\partial t}.$$

We compare z with $v(t) = -C_3(t+1)$ where C_3 is a positive constant satisfying

$$\text{(2.11)} \qquad C_3 > \max\left(|g|_0^\varrho, \, |h|_0^B\right), \quad C_3 > \left|k \frac{\partial f}{\partial t}\right|_0^\varrho.$$

Then

$$\frac{\partial v}{\partial t} < k(x, u) \, L_0 v + k(x, u) \frac{\partial f(x, t)}{\partial t},$$

and $v < z$ on $B + S_\varrho$. Using Lemma 1, we conclude that

$$\text{(2.12)} \qquad z(x, t) > -C_3(\varrho + 1) \text{ in } D_\varrho.$$

To find an upper bound on z, we compare z with $v=\varepsilon\varphi(x)$ where $\varphi(x)$ is defined in (2.3) and $\varepsilon>0$. By (2.8) and (A_7), z satisfies

$$\frac{\partial z}{\partial t} \leq k(x,u)\, L_0 z + C_4 z^2 \text{ in } D_\varrho, \tag{2.13}$$

where C_4 is a positive constant independent of u. On the other hand, if ε is sufficiently small, then

$$\frac{\partial v}{\partial t} \geq k(x,u)\, L_0 v + C_4 v^2 \text{ in } D_\varrho.$$

Since $v>0\geqq z$ on $B+S_\varrho$, we can use Lemma 1 and get

$$z(x,t) \leqq \varepsilon\varphi(x). \tag{2.14}$$

Having completed the proof that $\partial u/\partial t$ is *a priori* bounded, we apply an *a priori* estimate for elliptic equations due to CORDES [1] (see the remark following Theorem 2), which can be applied under our differentiability assumptions on L_0. We obtain an *a priori* bound on the Hölder coefficient of u with respect to x, the exponent being any positive number $\beta\leq 1$. Combining this bound with the bound on $\partial u/\partial t$, we get an *a priori* bound on $|u|_\beta^\varrho$ for any $0<\beta\leqq 1$. The proof of Theorem 3 can now be completed very similarly to the proof of Theorem 1.

Remark 1. The assumptions

$$\frac{\partial f}{\partial t} \leqq 0, \quad \frac{\partial g}{\partial t} \leqq 0, \quad L_0 h(x) - f(x,0) \leqq 0 \tag{2.15}$$

were used only in estimating $\partial u/\partial t$ from above. It is easy to see that the proof works well also if (2.15) is replaced by

$$\frac{\partial f}{\partial t} \leqq \varepsilon, \quad \frac{\partial g}{\partial t} \leqq \varepsilon, \quad L_0 h(x) - f(x,0) \leqq \varepsilon \qquad (\varepsilon>0) \tag{2.16}$$

provided ε is sufficiently small. If, however, ε is not small, the proof fails. Indeed, we show this by exhibiting a function z which satisfies

$$z_t \leqq z_{xx} + A z^2 \quad \text{in} \quad 0<x<1, \quad 0<t<4N \qquad (A>0,\ N>0)$$

$$z(x,0) = z(0,t) = z(1,t) = \frac{\mu}{N} \qquad (\mu>0)$$

and which tends to infinity as $x=\frac{1}{2}$, $t\to 4N$ provided $A\mu>8N+\frac{1}{4}$. The required function is

$$z(x,t) = \frac{\mu}{N - t\,x(1-x)}.$$

Remark 2. *Let the assumptions of Theorem 3 hold, and, in addition, assume that $c(x)\equiv 0$, $f(x,t)\to f(x)$, $g(x,t)\to g(x)$ uniformly in $x\in\overline{B}$ and $x\in\partial B$ respectively as $t\to\infty$. Let $g(x)>0$ on ∂B. Then $u(x,t)\to u^0(x)$ as $t\to\infty$ uniformly in $x\in\overline{B}$, and u^0 satisfies the elliptic system*

$$L_0 u^0(x) = f(x) \quad \text{for} \quad x\in B, \tag{2.17}$$

$$u^0(x) = g(x) \quad \text{for} \quad x\in\partial B. \tag{2.18}$$

Proof. The existence of u^0 is well known [14], [10]. Now let $w=u-u^0$, and consider the system which w satisfies:

$$\begin{aligned}
&\frac{\partial w}{\partial t} = k(x,u)\,L_0\,w + k(x,u)\,[f(x) - f(x,t)] \text{ in } D,\\
(2.19)\qquad &w(x,0) = h(x) - g(x) \text{ on } B,\\
&w(x,t) = g(x,t) - g(x) \text{ on } S.
\end{aligned}$$

From the part (i) of the proof of Theorem 3 we conclude, recalling that $c(x)=0$, that u is bounded in D from above and from below by positive constants. Hence, the differential equation in (2.18) is uniformly parabolic in D. We then can apply a theorem of [4] and conclude that $w(x,t)\to 0$ as $t\to\infty$, uniformly in $x\in\bar{B}$.

Remark 3. Theorem 3 and the proposition in the above remark can be extended to the case that $c=c(x,t)$ provided $\partial c/\partial t \leqq 0$ and (for the proposition to hold) provided $c\to 0$ sufficiently fast as $t\to\infty$.

If the boundary values h, g are not strictly positive but only nonnegative, then we can approximate h, g by positive functions h_m, g_m. The corresponding solutions u_m converge uniformly in each D_ϱ to a function u which may be considered as a "weak solution".

Remark 4. Consider the equation (0.3). In the physical problem arising from flow of gases, the boundary values are likely to be positive functions. We then may define $v=\psi(x,u)$ $(u>0)$ in D_ϱ with ϱ sufficiently small, and reduce (0.3) to (0.3'). Using the positive lower bound for solutions found in the proof of Theorem 3, we conclude that a similar bound exists for u. Hence the transformation $v=\psi(x,u)$ $(u>0)$ is valid in $D_{\varrho+\varepsilon}$ for some $\varepsilon>0$. Proceeding by a standard argument of continuation, we conclude that *the equation* (0.3) *is equivalent to the equation* (0.3'). We can therefore apply Theorem 3 to the system (0.3), (1.2). Note that the assumptions $k>0$, $\partial k/\partial v \geqq 0$ are equivalent to $\partial\psi/\partial u>0$, $\partial^2\psi/\partial u^2 \geqq 0$ and this is indeed satisfied in the special case of (0.2).

We conclude by mentioning that equation (0.3) has been studied by several authors in the special case $n=1$. The most complete treatment is in a recent work of Olainik, Kalashnikov & Yooie-Lynn [12], which contains references to earlier works. The methods of [12], however, are of very special nature.

Prepared under ONR Contract Nonr-222(62) (NR 041 214) with the University of California in Berkeley.

References

[1] Cordes, H. O.: Über die erste Randwertaufgabe bei quasilinearen Differentialgleichungen zweiter Ordnung in mehr als zwei Variablen. Math. Ann. **131**, 278–312 (1956).

[2] De Giorgi, E.: Sulla differenziabilità e l'analiticità delle estremali degle integrali multipli regolari. Memorie dell' Accademia delle Scienze di Torino, s. 3ª to. 3, Parte I, pp. 25–43 (1957).

[3] Friedman, A.: Boundary estimates for second order parabolic equations and their applications. J. Math. and Mech. **7**, 771–792 (1958).

[4] Friedman, A.: Convergence of solutions of parabolic equations to a steady state. J. Math. and Mech. **8**, 57–76 (1959).

[5] Ladyzhenskaya, O. A.: First boundary value problem for quasi-linear parabolic equations. Doklady Akad. Nauk SSSR (N.S.) **107**, 636—639 (1956).
[6] Morrey, C. B.: Second order elliptic equations in several variables and Hölder continuity. Technical Report No. 5, Univ. of Calif., Berkeley, March, 1959.
[7] Muskat, M.: The Flow of Homogeneous Fluids through Porous Media. Ann Arbor (Michigan) 1946.
[8] Nash, J.: Continuity of solutions of parabolic and elliptic equations. Amer. J. Math. **80**, 931—954 (1958).
[9] Nirenberg, L.: A strong maximum principle for parabolic equations. Comm. Pure Appl. Math. **6**, 167—177 (1953).
[10] Nirenberg, L.: Existence Theorems in Partial Differential Equations. New York University.
[11] Olainik, O. A., & T. D. Ventsel: First boundary problem and the Cauchy problem for quasi-linear equations of parabolic type. Mat. Sbornik **41**, (83) 105—128 (1957).
[12] Olainik, O. A., A. S. Kalashnikov & Chou Yooie-Lynn: The Cauchy problem and boundary problems for equations of the type of nonsteady filtrations. Izv. Akad. Nauk. SSSR, Sec. Mat. **22**, 667—704 (1958).
[13] Schauder, J.: Der Fixpunktsatz in Funktionalräumen. Studia Math. **2**, 171—180 (1930).
[14] Schauder, J.: Über lineare elliptische Differentialgleichungen zweiter Ordnung. Math. Z. **38**, 257—282 (1933/34).
[15] Ventsel, T. D.: First boundary problem and the Cauchy problem for quasi-linear parabolic equations with many space dimensions. Mat. Sbornik **41**, 499—520 (1957).
[16] Westphal, H.: Zur Abschätzung der Lösungen nichtlinearer parabolischer Differentialgleichungen. Math. Z. **51**, 690—695 (1949).

University of California
Berkeley, California
and
University of Minnesota
Minneapolis, Minnesota

(Received February 13, 1960)

Nouveau Fondement Fonctionnel de la Mécanique

Victor Vâlcovici

Mémoire transmis par C. Truesdell

Introduction

Lagrange dans sa célèbre «Mécanique analytique» a transformé la Mécanique newtonienne dans un chapitre de l'Analyse. Il pensait que la transformation serait complète sans aucun souvenir physique; il était même très fier d'avoir évité l'emploi de toute figure dans le texte [1].

La méthode de Lagrange était basée sur une nouvelle notion, celle du déplacement virtuel, ainsi que sur deux axiomes:

I. *Axiome de la libération* („Befreiungsprinzip" chez Hamel) *ou de l'existence* (∃). *Il existe toujours un système de forces — «les réactions» — qui ont la propriété de libérer le système de points, des obligations imposées par les liaisons.* On pourra représenter symboliquement, cet axiome par la relation suivante:

$$[(F)+(R)]\,(A) \simeq [(F)+(\mathscr{L})]\,(A);$$

ici (A) signifie le système de points considérés auquel on applique d'un côté, l'opérateur (F) de forces ainsi que celui des réactions (R) et de l'autre, l'opérateur (F) des forces associé à l'opérateur $(\mathscr{L})$ des liaisons, le signe $\simeq$ montrant que les deux opérations sont équivalentes.

II. *Axiome du travail virtuel des réactions* (U). *Le travail élémentaire correspondant à un déplacement virtuel quelconque du système est égal à zéro.* C'est grâce à cet axiome que l'on peut s'assurer l'unicité des réactions.

Moyennant la notion du déplacement virtuel et avec l'aide de deux axiomes on parvient à fournir à la Mécanique une position fonctionnelle susceptible d'être traitée par les méthodes de l'Analyse.

La Mécanique analytique ainsi conçue par Lagrange représente une élégante conquête des mathématiques.

Néanmoins quelques fissures ont apparu au cours des années dans cette merveilleuse construction — des questions ayant un caractère paradoxal qui troublent l'harmonie logique de l'oeuvre. Je cite notamment:

---

[1] [*1*] *Avertissement de la première édition.* «On ne trouvera point de Figures dans cet Ouvrage. Les méthodes que j'y expose ne demandent ni constructions, ni raisonnements géométriques ou méchaniques, mais seulement des opérations algébriques, assujetties à une marche régulière et uniforme. Ceux qui aiment l'Analyse, verront avec plaisir la Méchanique en devenir une nouvelle branche et me sauront gré d'en avoir étendu ainsi le domaine.»

Il est à peine nécessaire de mentionner que depuis, les idées concernant la Mécanique ont beaucoup changé: le support physique joue à l'heure actuelle un rôle dominant.

a) Différence de méthode exigée dans le traitement mathématique des diverses catégories de liaisons (holonomes, non holonomes de première et de seconde classe)[2].

b) Distinction entre les déplacements virtuels figurant d'un côté, dans l'équation fondamentale de la Mécanique (Principe de D'ALEMBERT-LAGRANGE) et de l'autre, dans les principes variationnels[3].

c) Déviations par rapport à l'axiome du travail des réactions[4].

Je me propose de substituer à la notion de déplacement virtuel ainsi qu'aux deux axiomes formulés ci-dessus, d'autres constructions qui soient exemptes des défauts signalés.

I. Le Fondement Fonctionnel

§ 1. Variation du mouvement

Supposons que les positions des n points matériels A_i $(i=1, 2, \ldots, n)$ ayant les masses m_i et constituant le système (A) soient déterminées par les vecteurs x_i dans l'espace euclidien E_3 et leur mouvement par des relations de la forme

$$x_i = x_i(t), \tag{1}$$

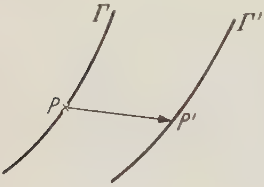

Fig. 1

$x_i(t)$ étant des fonctions dérivables du temps t, (classe C_2), définies dans l'intervalle

$$t = [t_0, t_1], \qquad t_0 < t_1. \tag{2}$$

Soit Γ Fig. 1, la trajectoire d'un point quelconque A du système, ayant la masse m, P étant sa position au moment t, et désignons par δx un vecteur infinitésimal, fonction dérivable de t dans tout l'intervalle (2) telle que sa dérivée soit également infinitésimale du même ordre de grandeur. Le point P' ayant le vecteur de position $x' = x(t) + \delta x$ décrira la courbe Γ' lorsque P décrira la courbe Γ; la courbe Γ' sera voisine de Γ. Imaginons un mobile fictif A' sur Γ' qui se trouve en P au moment t' $(=t+\delta t)$, où nous avons désigné par δt une fonction dérivable de t, définie dans l'intervalle (2) et ayant l'ordre de grandeur de δx, ainsi que sa dérivée. Le mouvement de A' sur Γ' sera une *variation* du mouvement de A sur Γ. Les deux mouvements considérés — celui de A sur Γ et celui de A' sur Γ' — seront *synchrones* seulement dans le cas $\delta t \equiv 0$ dans tout l'intervalle (2); dans le cas contraire ils seront *non synchrones*.

§ 2. Variation d'une fonctionnelle

Nous supposerons que les fonctions δx et δt de t sont arbitraires, en faisant toutefois la réserve qu'elles s'annulent aux deux bouts de l'intervalle (2).

Le mouvement le long de Γ ainsi que le long de Γ' est réalisé à l'aide de l'opérateur d,

$$dx, \quad dt, \quad dx', \quad dt', \quad \text{etc.},$$

tandis que la variation de Γ vers Γ' sera représentée à l'aide de l'opérateur δ,

$$\delta x, \quad \delta t \quad \text{etc.}$$

[2] [8] p. 507.
[3] [8] p. 523.
[4] [3] p. 272, [4].

On prendra, comme d'habitude,

$$\delta\, dt = d\, \delta t, \tag{3}$$

ce qui est juste si l'on admet que l'opérateur δ, appliqué à un point de Γ mène toujours à un point de Γ'.

Explicitons maintenant l'expression

$$\delta(f\, dt),$$

où l'on a désigné par f une fonction $f(x, \dot{x}, t)$ définie le long de Γ dans tout l'intervalle (2) et possédant les dérivées partielles des deux premiers ordres par rapport aux arguments x, $\dot{x}$, t. Ayant égard à (3) on pourra écrire

$$\delta(f\, dt) = \left(\delta f + f \frac{d}{dt}\delta t\right) dt. \tag{4}$$

Mais l'on a

$$\delta f = f_x\, \delta x + f_{\dot{x}}\, \delta \dot{x} + f_t\, \delta t, \tag{5}$$

avec la notation

$$f_x = \nabla_x f, \qquad f_{\dot{x}} = \nabla_{\dot{x}} f. \tag{6}$$

$\delta \dot{x}$ signifie la variation de la vitesse $v\,(=\dot{x})$ lorsqu'on passe du point P au point P'. On aura donc

$$\delta \dot{x} = \frac{dx'}{dt'} - \frac{dx}{dt} = \frac{d(x+\delta x)}{d(t+\delta t)} - \frac{dx}{dt} = \left(\frac{dx}{dt} + \frac{d}{dt}\delta x\right)\left(1 + \frac{d}{dt}\delta t\right)^{-1} - \frac{dx}{dt},$$

ou bien, en négligeant les quantités d'un ordre supérieur au premier,

$$\delta \dot{x} = \frac{d}{dt}\delta x - \dot{x}\frac{d}{dt}\delta t. \tag{7}$$

Dans le cas particulier de la variation synchrone ($\delta t = 0$) cette relation deviendra

$$\delta_0 \dot{x} = \frac{d}{dt}\delta x, \tag{8}$$

en notant avec $\delta_0 \dot{x}$ la variation correspondant à la vitesse $\dot{x}\,(=v)$ dans ce cas particulier. Donc la formule (7) pourra être mise sous la forme

$$\delta \dot{x}\,(=\delta v) = \delta_0 \dot{x} - \dot{x}\frac{d}{dt}\delta t. \tag{9}$$

Moyennant cette formule on pourra déterminer la variation de l'énergie cinétique T du système (A) de points,

$$T = \tfrac{1}{2}\sum m\, v^2, \tag{10}$$

la somme étant étendue aux n points du système. On aura

$$\delta T = \sum m\, v\, \delta v, \tag{11}$$

ou bien, avec (9),

$$\delta T = \sum m\, v\, \delta_0 \dot{x} - \sum m\, v^2 \frac{d}{dt}\delta t.$$

Le premier terme de droite peut être interprété comme représentant la variation $\delta_0 T$ de l'énergie cinétique T dans le cas des mouvements variés synchrones,

$$\sum m\, v\, \delta_0 \dot{x} = \delta_0 T. \tag{12}$$

De cette manière l'expression de δT prendra la forme

$$\delta T = \delta_0 T - 2T \frac{d}{dt}\,\delta t. \tag{13}$$

En revenant à la formule (7) nous la mettrons aussi sous la forme

$$\delta \dot{x}\,(=\delta v) = \frac{d}{dt}\,(\delta x - \dot{x}\,\delta t) + \ddot{x}\,\delta t, \tag{14}$$

où l'on a mis en évidence l'opérateur

$$\Delta = \delta - \delta t \frac{d}{dt} \tag{15}$$

appliqué au vecteur x,

$$\Delta x = \delta x - \dot{x}\,\delta t.$$

Cet opérateur trouvera un usage fréquent dans ce qui suit. Par suite, la relation (14) deviendra

$$\delta \dot{x}\,(=\delta v) = \frac{d}{dt}\,\Delta x + \ddot{x}\,\delta t \tag{16}$$

et la formule (4) avec (5) et (16),

$$\delta(f\,dt) = f_x\,\delta x\,dt + f_{\dot{x}}\,d\,\Delta x + (f_{\dot{x}}\ddot{x} + f_t)\,dt\,\delta t + f\,d\,\delta t,$$

ou bien

$$= d\,(f_{\dot{x}}\,\Delta x + f\,\delta t) - \Delta x\,d f_{\dot{x}} - d f\,\delta t + f_x\,\delta x\,dt + (f_{\dot{x}}\ddot{x} + f_t)\,dt\,\delta t.$$

Faisons usage des formules évidentes

$$d f = (f_x \dot{x} + f_{\dot{x}}\ddot{x} + f_t)\,dt,$$

$$-f_x \dot{x}\,\delta t\,dt + f_x\,\delta x\,dt = f_x\,\Delta x\,dt.$$

On obtiendra

$$\delta(f\,dt) = d\,(f_{\dot{x}}\,\Delta x + f\,\delta t) + \left(f_x - \frac{d}{dt}\,f_{\dot{x}}\right)\Delta x\,dt. \tag{17}$$

Dans cette formule on remarque aussi l'opérateur eulérien

$$\vartheta = \frac{d}{dx} - \frac{d}{dt}\,\frac{\partial}{\partial \dot{x}}$$

appliqué à la fonction f; de cette façon on pourra mettre la formule (17) sous la forme

$$\delta(f\,dt) = d\,(f_{\dot{x}}\,\Delta x + f\,\delta t) + \vartheta f\,\Delta x\,dt. \tag{18}$$

§ 3. Coordonnées généralisées. Espace de Lagrange

En passant à l'espace de LAGRANGE le mouvement du point représentant le système (A) sera donné par les équations

$$q^i = q^i(t), \qquad i = 1, 2, \ldots, s,$$

$q^i(t)$ étant des fonctions de classe C_2 dans l'intervalle (2). Une variation quelconque de ce mouvement sera déterminée par les équations

$$q^i = q'^i(t), \quad q'^i(t) = q^i(t) + \delta q^i, \quad t' = t + \delta t, \tag{19}$$

où δq^i, δt signifient des fonctions infinitésimales de t, ayant comme dérivées des fonctions également infinitésimales de t, de l'ordre de grandeur de δq^i et δt. En désignant par C la trajectoire dans l'espace de Lagrange (q) et par Q la position du point représentatif au moment t quelconque dans l'intervalle (2) on obtiendra une variation de ce mouvement sur une courbe voisine C' à l'aide des fonctions infinitésimales δq^i, δt.

En utilisant les procédés du § 2 nous obtiendrons les formules

$$\delta \dot{q} = \frac{d}{dt} \Delta q + \ddot{q}\, \delta t, \tag{20}$$

$$\delta (f\, dt) = d(f_{\dot{q}} \Delta q + f\, dt) + \vartheta f \Delta q\, dt \tag{21}$$

à la place des formules (16) et (18), f étant ici une fonction de $q, \dot{q}, t$ et l'opérateur eulérien ϑ ayant l'expression

$$\vartheta = \frac{\partial}{\partial q} - \frac{d}{dt} \frac{\partial}{\partial \dot{q}}.$$

§ 4. *Variation d'une intégrale curviligne*

Nous nous proposons dans ce qui suit d'étudier la genèse des principes variationnels de la Mécanique en partant d'un problème de calcul fonctionnel.

Considérons l'intégrale

$$I(\Gamma) = \int_\Gamma f(x, \dot{x}, t)\, dt \tag{22}$$

dans l'espace réel. Le passage de Γ à Γ' nous donnera

$$\delta I = I(\Gamma') - I(\Gamma) = \int_{\Gamma'} f\, dt - \int_\Gamma f\, dt,$$

ou bien, en faisant la substitution (19) dans $I(\Gamma')$,

$$\delta I = \int_\Gamma \delta (f\, dt). \tag{23}$$

En utilisant la formule (18) on obtiendra

$$\delta I = \int_\Gamma \vartheta f \Delta x\, dt, \tag{24}$$

car Δx et δt s'annulent aux deux bouts de l'intervalle (2).

Nous aurons également

$$\delta I = \int_C \vartheta f \Delta q\, dt \tag{25}$$

en coordonnées généralisées. La condition pour que la fonctionnelle I soit stationnaire sera donc

$$\int \vartheta f \Delta x\, dt = 0, \tag{26}$$

respectivement

$$\int \vartheta f \Delta q\, dt = 0. \tag{27}$$

§ 5. *Principe d'Hamilton-Ostrogradski*

Le principe variationnel qui a rendu les plus grands services à la Mécanique est, sans aucun doute, celui D'HAMILTON-OSTROGRADSKI. Dans sa forme classique ce principe suppose que les mouvements variés admissibles sont synchrones.

Le mathématicien italien E. STORCHI [5] est le premier en date (1955) qui ait montré que dans le cas des forces conservatives on peut imaginer certains mouvements variés admissibles, non synchrones, par rapport auxquels le principe d'Hamilton-Ostrogradski soit valable.

Les cas des forces quasi-conservatives. Nous avons donné ce nom aux forces admettant une fonction des forces $U(x_1, \ldots, x_n, t)$ qui dépende explicitement du temps t aussi. On aura dans ce cas pour les forces F_i $(i=1, 2, \ldots, n)$ appliquées aux n points A_i du système (A), la formule

$$F_i = U_{x^i}$$

signifiant au fond

$$F_i = \nabla_i U.$$

Le travail élémentaire $\delta\mathscr{L}$ des forces F, correspondant aux variations δx du système aura l'expression

$$\delta\mathscr{L} = \sum F\,\delta x = \Sigma U_x\,\delta x. \tag{28}$$

Mais l'on a

$$\delta U = \sum U_x\,\delta x + U_t\,\delta t, \tag{29}$$

de sorte que la formule (28) du travail deviendra

$$\delta\mathscr{L} = \delta U - U_t\,\delta t. \tag{30}$$

Dans le cas des variations synchrones $(\delta t \equiv 0)$ δU se réduit à $\delta_0 U$ et l'on pourra écrire par conséquent

$$\delta U = \delta_0 U + U_t\,\delta t. \tag{31}$$

Tâchons d'établir une connexion entre la formule générale (25) obtenue ci-dessus et le Principe d'Hamilton-Ostrogradski en supposant que les forces sont quasi-conservatives. A cet effet nous ferons $f = L$ dans (22), L étant la fonction de Lagrange

$$L = T + U. \tag{32}$$

La formule (25) deviendra

$$\delta I = \int \vartheta L\,\Delta x\,dt. \tag{33}$$

Afin de trouver plus rapidement l'expression explicite de cette formule remplaçons f dans (4) f par T et par U successivement. Nous aurons

$$\delta(T\,dt) = \left(\delta T + T\frac{d}{dt}\,\delta t\right)dt, \tag{34}$$

ou bien, avec (13),

$$\delta(T\,dt) = \left(\delta_0 T - T\frac{d}{dt}\,\delta t\right)dt, \tag{34'}$$

ainsi que

$$\delta(U\,dt) = \left(\delta U + U\frac{d}{dt}\,\delta t\right)dt, \tag{35}$$

ou bien

$$\delta(U\,dt) = \left(\delta_0 U + U_t\,\delta t + U\frac{d}{dt}\,\delta t\right)dt \tag{35'}$$

eu égard aux formules

$$\delta U = U_x\,\delta x + U_t\,\delta t, \qquad \delta_0 U = U_x\,\delta x.$$

---

[5] [12].

L'addition des formules (34′) et (35′) donnera

$$\delta(L\,dt) = \left(\delta_0 L - E\,\frac{d}{dt}\,\delta t + U_t\,\delta t\right) dt, \tag{36}$$

où l'on a mis

$$E = T - U. \tag{37}$$

Par suite, la relation (33) deviendra

$$\delta I = \int \delta_0 L\,dt - \int \left(E\,\frac{d}{dt}\,\delta t - U_t\,\delta t\right) dt. \tag{38}$$

La condition $\delta I = 0$ pour que la fonctionnelle I soit stationnaire s'écrit donc

$$\int \delta_0 L\,dt - \int \left(E\,\frac{d}{dt}\,\delta t - U_t\,\delta t\right) dt = 0. \tag{39}$$

Cette formule représente la forme généralisée du principe d'Hamilton-Ostrogradski; car elle prendra la forme classique de ce principe

$$\int \delta_0 L\,dt = 0 \tag{40}$$

et partant, l'expression

$$\int L\,dt$$

aura une valeur stationnaire si l'on peut disposer des variations δx, δt afin que la condition

$$\int \left(E\,\frac{d}{dt}\,\delta t - U_t\,\delta t\right) dt = 0 \tag{41}$$

soit remplie. Les variations non synchrones correspondantes représentent les mouvements variés admissibles pour le principe d'Hamilton-Ostrogradski, dont les mouvements variés synchrones de la forme classique en forment un cas particular.

Cette importante observation a été faite pour la première fois par E. Storchi[6]. Celui-ci a limité l'observation au cas particular des systèmes conservatifs ($U_t \equiv 0$); pourtant elle peut être étendue aux systèmes quasi-conservatifs, comme on a vu ci-dessus, avec la condition (41). Par ailleurs, on peut en déduire celle donnée par Storchi en y faisant $U_t \equiv 0$.

§ 6. *Principe de moindre action*

En prenant $f \equiv 2T$ dans (26) on arrive au principe de moindre action (Maupertuis) sous la forme

$$\int 2T\,dt = \text{stationnaire.} \tag{42}$$

Laissant de côté le facteur 2 et posant

$$I = \int T\,dt \tag{43}$$

on arrive avec (34) à la condition

$$\int \left(\delta T + T\,\frac{d}{dt}\,\delta t\right) dt = 0. \tag{44}$$

[6] [*12*].

Donc il suffit de déterminer une fonction δt satisfaisant à la relation (44) pour que le principe (42) soit valable. Moyennant l'identité

$$\frac{d}{dt}(T\,\delta t) = T\frac{d}{dt}\,\delta t + \dot{T}\,\delta t$$

nous ferons une intégration partielle en (44) du dernier terme. On aura

$$\int (\delta T - \dot{T}\,\delta t)\,dt = 0, \tag{45}$$

vu que δt est nul aux deux bouts de l'intervalle (2). Mais si le système admet le théorème de l'énergie sous la forme

$$\dot{T} = \sum F\,v$$

la condition (45) deviendra

$$\int (\delta T - \sum F\,v\,\delta t)\,dt = 0. \tag{46}$$

Il s'ensuit que le principe de moindre action sera valable pour tous les mouvements variés satisfaisant à la condition (46), en particulier, pour les mouvements variés satisfaisant à la condition

$$\delta T = \sum F\,v\,dt. \tag{47}$$

§ 7. *Autres principes variationnels*

La relation (26), respectivement (27), pourrait engendrer d'autres principes similaires, en dehors des deux classiques bien connus, en prenant pour f d'autres fonctions en dehors de L et $2T$; par exemple, $f = U$ ou $f = E$. Dans le premier cas ($f = U$) la formule (35) donnerait

$$\int \left(\delta U + U\frac{d}{dt}\,\delta t\right) dt = 0, \tag{48}$$

une condition pour δt afin d'obtenir le principe

$$\int U\,dt = \text{stationaire.}$$

Dans le second ($f \equiv E$) la condition correspondante pour δt serait

$$\int \left(\delta E + E\frac{d}{dt}\,\delta t\right) dt = 0,$$

ce qui conduirait au principe

$$\int E\,dt = \text{stationnaire.}$$

Si l'on prenait d'une manière plus générale

$$f = \alpha_1 T + \alpha_2 U,$$

α_1, α_2 étant deux constantes arbitraires on pourrait obtenir une infinité de principes variationnels[7].

[7] Dans [*16*] j'ai exposé aussi d'autres méthodes pour générer des principes nouveaux.

II. Equation Fondamentale de la Mécanique

§ 8. *Condition de liaison*

On entend d'habitude par *condition de liaison* une entrave de nature géométrique ou cinématique dans le mouvement du système (A).

Une analyse plus détaillée du mouvement[8] met en évidence l'importance du vecteur dx $(=v\,dt)$ qui représente le déplacement réel du point considéré dans l'intervalle dt de temps. L'élément dx est pour ainsi dire la «cellule constitutive» du mouvement «global» du point. Il s'ensuit qu'une «condition de liaison» imposée au point A devrait attaquer en premier lieu la liberté de l'élément dx; subséquemment celle des éléments $x, \dot{x}, \ddot{x}, t$. Néanmoins dans la forme classique de la Mécanique on met l'habitude les conditions de liaison sous la forme[9]

$$\varphi(x, \dot{x}, \ddot{x}, t) = 0, \tag{49}$$

ce qui n'est pas normal. Il serait plus naturel de les représenter par des relations

$$\varphi(dx, dt) = 0 \tag{50}$$

qui lient entre eux les paramètres essentiels du mouvement, dx et dt, les autres éléments du mouvement $x, \dot{x}, \ddot{x}, t$ pouvant y entrer aussi. Mais la nature infinitésimale de l'équation fondamentale, obtenue à l'aide du principe de d'Alembert-Lagrange impose à la relation (50) une forme linéaire

$$\alpha\, dx + \beta\, dt = 0, \tag{51}$$

le vecteur α et le scalaire β étant des fonctions de $x, \dot{x}, \ddot{x}, t$ et $\alpha\, dx$ un produit scalaire. Pour éviter toute erreur il faut accentuer cette propriété de α et β qu'ils ne sont pas astreints à dépendre seulement de x et t, comme on suppose dans la forme classique de la Mécanique, mais des vecteurs $\dot{x}, \ddot{x}$ aussi par leurs composantes[10]. Avec ce complètement on peut affirmer que les formes linéaires (51) comprennent les conditions les plus générales (holonomes ou non holonomes) qui pourraient intervenir dans la réalité[10].

En employant des indices, on aura donc pour les m liaisons linéaires

$$\alpha_i^j\, dx_i + \beta^j\, dt = 0, \qquad i = 1, 2, \ldots, n; \quad j = 1, 2, \ldots, m; \tag{52}$$

nous avons supprimé le signe $\sum_i$ qui devrait précéder le premier terme — ce que nous allons faire en général, dans la suite. Comme conséquence de la forme (52) adoptée pour les relations de liaison, on ne sera plus obligé d'employer des méthodes différentes pour traiter les liaisons holonomes et celles non holonomes; elles seront traitées toutes par une méthode unitaire.

Nous supposerons en outre que les m relations scalaires (52) sont distinctes[10]. Elles fourniront donc au problème analytique m relations distinctes entre les éléments $x, \dot{x}, \ddot{x}, t$ du mouvement, car en y remplaçant dx par $\dot{x}\, dt$ on aura

$$\alpha_i^j\, \dot{x}_i + \beta^j = 0. \tag{53}$$

---

[8] [*16*].

[9] En nous limitant aux liaisons bilatérales; d'ailleurs les liaisons unilatérales (représentées par des égalités et des inégalités) se réduisent à celles-là.

[10] [*16*].

§ 9. *Equation fondamentale*

Le principe de d'Alembert-Lagrange qui coïncide avec le principe que Lagrange appelle des «vitesse virtuelles», appliqué à la Dynamique, nous fournit *l'équation fondamentale de la Mécanique*

$$(-m_i \ddot{x}_i + F_i)\,\delta x_i = 0, \tag{54}$$

où nous avons supprimé, comme ci-dessus, le signe $\sum\limits_i$. Nous écrirons cette équation sous la forme abrégée

$$\Phi_i\,\delta x_i = 0, \tag{55}$$

avec la notation

$$\Phi_i = -m_i \ddot{x}_i + F_i. \tag{56}$$

Nous nommerons le vecteur Φ «vecteur d'Alembert».

Maintenant surgit une question qui présente une grande importance pour la résolution du problème: quelle est la répercussion que les conditions de liaison (51) doivent avoir sur les déplacements virtuels?

La forme classique de la Mécanique a donné à cette question la réponse connue: les déplacements virtuels δx_i devraient satisfaire aux relations que l'on obtiendrait de (52) en y faisant $d x_i = \delta x_i$ et en y considérant t comme constant, c.à.d en y faisant $dt = 0$. Il s'ensuit que les déplacements virtuels δx_i *admissibles* pour l'équation fondamentale (55) devraient satisfaire à la condition

$$\alpha_i^j\,\delta x_i = 0. \tag{57}$$

La méthode des multiplicateurs de Lagrange livrera dans ce cas la solution du problème sous la forme du système d'équations

$$\Phi_i + \sum_j \lambda_j \alpha_i^j = 0 \tag{58}$$

auquel on adjoint le système (53). Le système est donc déterminé, le nombre des équations étant égal à celui des inconnues x_i, λ_j.

La justification physique du passage des relations (51) aux conditions (57) imposées aux déplacements virtuels δx_i admissibles est basée sur la propriété axiomatique des réactions R_i,

$$R_i\,\delta x_i = 0, \tag{59}$$

et celle-ci sur le «fait avéré» — reconnu comme tel — que le mouvement du point obligé à se mouvoir sur une surface immobile et parfaitement lisse (dépourvue de frottement) subit de la part de la surface une réaction normale.

Il est évident que dans ce cas particulier la relation (59) est satisfaite si les déplacements virtuels admissibles δx_i sont limitées aux déplacements sur la surface même. Or cette dernière condition s'exprime par une relation de la forme (57) et celle-ci provient de (52) en y faisant $dt = 0$, $d x_i = \delta x_i$.

La relation (59) apparait donc comme une extension naturelle du phénomène correspondant à la surface immobile et dépourvue de frottement.

§ 10. *Désaccord entre l'équation fondamentale et les fondements fonctionnels de la Mécanique*

L'examen mathématique du fondement fonctionnel de la Mécanique conduit, comme on a vu ci-dessus, à la relation (24) qui contient Δx à la place où devrait figurer δx. Cette substitution

$$\Delta x \to \delta x \tag{60}$$

rend inutilisable l'élégant procédé des multiplicateurs si le problème comporte des relations de liaisons ayant la forme (52) avec $\beta^j \not\equiv 0$ (système rhéonome). Il y a donc un désaccord entre l'équation fondamentale, d'un côté, et la base fonctionnelle de la Mécanique, de l'autre, non seulement à cause de cette interdiction de méthode; car on constate aussi que la division infinitésimale du problème de la Mécanique emploie des unités différentes selon que cette division est réalisée moyennant le principe des déplacements virtuels ou bien moyennant l'ensemble des mouvements variés.

§ 11. Variations virtuelles (possibles)

Afin de supprimer le désaccord mentionné ci-dessus nous allons essayer de remplacer la notion classique de «déplacement virtuel» δx par une autre plus convenable. Nous avons été conduits à cette idée aussi par la mauvaise réputation dont jouit la notion du déplacement virtuel; car cette notion était depuis longtemps reconnue comme une notion artificielle, presque inadmissible [11]. Même GAUSS en s'y rapportant dit que cette notion ne possède pas de crédit pour pouvoir être recommandée comme plausible [12].

D'ailleurs on avait essayé même avant nous de la modifier. Le mathématicien italien LUIGI CASTOLDI a proposé en 1949 [13] une nouvelle définition pour le déplacement virtuel, en utilisant les «variations» bien connus de JOURDAIN. Mais malgré sa complication cette nouvelle notion ne parvient pas à mettre d'accord les deux quantités fondamentales, $\varDelta x$ et δx.

Le moyen le plus simple pour y arriver serait de faire la substitution (60), voilà comment.

Nous considérons δx et δt comme des variations arbitraires des vecteurs x et du temps t. Elles seront *possibles* (ou *virtuelles*) [14] seulement si elles satisfont aux relations

$$\alpha_i^j \, \delta x_i + \beta^j \, dt = 0 \tag{61}$$

analogues aux conditions de liaisons (52). Parmi les variations virtuelles il y aura évidemment, le déplacement réel du système — c.à.d. dx, dt — ainsi que le déplacement virtuel de la Mécanique classique pour $\delta t \equiv 0$. Il est clair que les «variations virtuelles» ont un degré de liberté en plus, si on les compare aux déplacements virtuels classiques satisfaisant à la condition (57), dû à l'apparition de l'élément arbitraire δt dans leur composition.

Pour distinguer les déplacements virtuels de la Mécanique classique ($\delta t \equiv 0$), des variations arbitraires δx nous allons les désigner par $\delta_0 x$.

Maintenant nous allons donner aux relations (61) une autre forme afin de rapprocher les «variations virtuelles», des quantités $\varDelta x$, en remplaçant β^j par l'expression

$$\beta^j = -\alpha_i^j \, \dot{x}_i$$

---

[11] [5] p. 49.

[12] [2] „Der eigentümliche Charakter des Prinzips der virtuellen Geschwindigkeiten besteht darin, daß es eine allgemeine Formel zur Auflösung aller statischen Aufgaben, und so der Stellvertreter aller anderen Prinzipe ist, ohne jedoch das Creditiv dazu so unmittelbar aufzuweisen, daß es sich, so wie es nur ausgesprochen wird, schon von selbst als plausibel empföhle."

[13] [10], [11].

[14] En langue russe on traduit «déplacement virtuel» par vozmojnoe pere mechnie c.à.d. «déplacement possible».

tirée de (53). Par suite de cette substitution la relation (61) deviendra

$$\alpha_i^j \delta x_i - \alpha_i^j \dot{x}_i \delta t = 0,$$

ou bien,

$$\alpha_i^j \Delta x_i = 0, \tag{62}$$

en utilisant l'opérateur Δ défini par la relation (15).

Cette égalité nous incite à essayer la substitution de Δx partout à la place du déplacement virtuel δx, comme cela se voit au passage (57)→(62); par conséquent, à essayer de remplacer la notion de «déplacement virtuel δx» par la notion de «variation virtuelle Δx».

On serait arrivé de cette façon à la substitution (60) que nous avions en vue.

L'accord entre les deux aspects de la Mécanique, — celui imposé par l'équation fondamentale et celui des principes variationnels —, *est ainsi réalisé* et par suite, la méthode des multiplicateurs sera également applicable dans le cas des principes variationnels aussi. Quant à l'équation fondamentale elle prendra la forme

$$\Phi_i \Delta x_i = 0, \tag{63}$$

au lieu de (55); associée aux relations (62) elle menera aux mêmes équations (58) de sorte que le système des équations conservera la forme du cas classique. *La substitution* (60) *ne change donc rien à la solution*, tout en réalisant l'accord entre la forme de l'équation fondamentale et celle du problème fonctionnel.

Il nous reste encore une chose capitale à accomplir. Il faudra justifier, au point de vue physique, la substitution (60) dans l'équation fondamentale, c'est-à-dire le passage (55)→(63). A cet effet il sera nécessaire de reprendre les deux axiomes concernant les réactions. Nous maintiendrons invariable le premier axiome ($\exists$), de la libération. Quant au second ($\cup$) il subira une certaine transformation de contenu.

Commençons pa la notion du «*travail virtuel des réactions*» qui aura maintenant une nouvelle définition conformément à l'élément Δx remplaçant le déplacement virtuel δx. Nous appellerons *travail élémentaire des réactions correspondant à une variation virtuelle δx, δt l'expression $R_i \Delta x_i$.*

Comparée avec l'expression classique $R_i \delta x_i$ du travail, cette nouvelle expression $R_i \Delta x_i$ possède un terme en plus dû au déplacement additionnel $\delta_a x_i$ $(= -\dot{x}_i \delta t)$ du point considéré avec sa vitesse réelle prise en sens opposé $(-\dot{x}_i)$ et pendant l'intervalle de temps δt. Donc le travail sera égal à $R_i(\delta x_i + \delta_a x_i)$ ce qui signifie que le travail virtuel dans sa nouvelle acception tient compte à la fois des deux sortes de paramètres: spatial (δx_i) et temporel (δt).

Avec cette modification de la notion du *travail virtuel* le second axiome peut être maintenu dans la forme que nous lui avions donnée au commencement:

II^a. *Le travail virtuel correspondant à une variation virtuelle Δx_i quelconque du système est égale à zéro.*

Nous aurons donc

$$R_i \Delta x_i = 0 \tag{64}$$

à la place de (59). Mais la réaction R_i est égale au vecteur Φ_i, donc

$$\Phi_i \Delta x_i = 0. \tag{65}$$

Ce sera la nouvelle forme de *l'équation fondamentale.*

Il faut pourtant se rappeler que la relation (59) — remplacée ici par (64) — était considérée comme possédant un puissant appui physique dans le fait que la réaction d'une surface immobile aura la direction de la normale, s'il n'y pas de frottement — phénomène considéré comme un cas particulier du cas général (rhéonome). Mais ce phénomène pourra continuer à jouer le rôle de puissant appui physique pour la nouvelle relation (64) aussi, car il peut être considéré également comme un cas particulier de la relation (64); donc, cette dernière — la relation (64) — pourra toujours apparaître comme une extension — plus évoluée — de celle qui correspond au cas scléronome de la surface immobile.

Conclusions

On voit de ce qui précède que l'élégante construction de la Mécanique analytique reste en vigueur telle qu'elle a été imaginée par LAGRANGE, dans les grandes lignes de son contour.

Néanmoins, afin d'obvier à certaines aspérités qui apparaissent au cours de son développement il est de rigueur d'entreprendre une révision de sa structure axiomatique qui est à la base de cette Mécanique devenue classique; en d'autres termes, les fondements même de la discipline devraient être de nouveau examinés.

Dans ce but nous avons commencé par établir une base fonctionnelle très large, au service des principes variationnels de la Mécanique, en leur fournissant ainsi un mécanisme commode et général de déduction.

On a vu de cette manière que les deux principes classiques — «D'HAMILTON-OSTROGRADSKI» et de «moindre action (MAUPERTUIS)» — peuvent être étendus à des mouvements variés admissibles constituant un ensemble bien plus puissant que son correspondant de la Mécanique classique.

Nous avons montré à cette occasion que le même mécanisme est en état d'engendrer une infinité d'autres principes variationnels en dehors des deux classiques bien connus.

Cet examen fonctionnel nous a conduit d'une manière naturelle à un nouvel élément différential, Δx, qui figure même dans la structure organique de la fonctionnelle.

L'apparition de Δx au sein de la fonctionnelle trouble profondément le maniement des principes variationnels au cas où il existe des liaisons; car l'élément différential δx employé dans le principe de d'Alembert-Lagrange comme déplacement virtuel ne s'harmonise pas avec l'élément Δx de la fonctionnelle. Pour éviter l'état déséquilibré qui en résulte nous avons été amenés à remplacer la «déplacement virtuel» de la Mécanique classique par la «variation virtuelle» et partant, δx par Δx.

Ce nouvel élément, Δx, implique un élargissement de l'ancien, δx, ayant un degré de liberté en plus et pouvant en conséquence, satisfaire les exigences logiques d'une manière plus convenable. L'opérateur Δ qui définit le nouvel élément, inclut aussi le temps, de sorte que l'élément lui-même dépendra également du temps t, non seulement de x. Le temps constant qui était imposé dans la Mécanique classique pour arriver à la définition du déplacement virtuel n'est plus ici nécessaire pour la «variation virtuelle» — et celle-ci remplace le déplacement virtuel dans toutes ses fonctions.

Par ailleurs, la solution de l'équation fondamentale — c.a.d. celle qui représente le mouvement du système — reste invariable: elle reproduit exactement la valeur fournie par la Mécanique classique, comme d'ailleurs la valeur des réactions aussi.

Mais afin de réaliser cet échange entre δx et Δx on a dû modifier aussi l'axiome du travail virtuel des réactions en le mettant d'accord avec la substitution (60).

Ces modifications qui réparent les fissures signalées au début de ce travail ont été accomplies sans violer aucunement l'allure physique des phénomènes. L'ensemble de toutes les notions dans leur nouvelle forme nous apparait comme une extension naturelle des notions classiques — extension qui amène avec elle un complètement de méthode ainsi qu'un complément de pouvoir opérateur.

Bibliographie

[1] Lagrange, J.-L.: Mécanique analytique 3e éd. revue par J. Bertrand, Paris 1853.
[2] Gauss, C. F.: Über ein neues allgemeines Grundgesetz der Mechanik. Crelle J. **IV** (1829) ou Werke, V, p. 25. 1877.
[3] Appell, P.: Rendiconti Palermo **33**, 259—267 (1912) ou Mémorial des Sci. math. **1** (1925).
[4] Delassus, E.: Ann. Ec. N. Sup. 1925.
[5] Nordheim, L.: Handbuch der Physik (H. Geiger u. K. Scheel), p. 43—90. Berlin: Springer 1927.
[6] Hamel, G.: Math. Ann. **92**, 33—41 (1924).
[7] Hamel, G.: Math. Ann. **111**, 94—97 (1935).
[8] Hamel, G.: Theoretische Mechanik. Berlin-Göttingen-Heidelberg: Springer 1949.
[9] Kerner, M.: Prace **38**, 1—21 (1931).
[10] Castoldi, L.: Rend. Ist. Lombardo **79**, 1—18 (1945/46).
[11] Castoldi, L.: Rend. Lincei **3**, 329—333 (1947).
[12] Storchi, E.: Rend Lincei **18**, 162—167 (1955).
[13] Vâlcovici, V.: Rend. Lincei **19**, 441—448 (1955).
[14] Vâlcovici, V.: C. R. Acad. Sci. Paris **243**, 1012—1014, 1096—1098 (1956).
[15] Vâlcovici, V.: Ber. sächs. Akad. Wiss., Leipzig **102**, 1—39 (1958).
[16] Vâlcovici, V.: Rev. de Méc. appl. **3**, 365—371 (1958).
[17] Vâlcovici, V.: Rev. de Math. pures et appl. **3**, 191—206 (1958).

Institut mathématique
de l'Académie de la République
Populaire Roumaine
Bucarest

(Manuscrit reçu le 4 janvier 1960)

Further Symmetry Relations for Transversely Isotropic Materials

J. E. ADKINS

Communicated by A. E. GREEN

1. Introduction

In earlier papers (ADKINS 1958, 1960) forms of the stress deformation relations have been investigated for a class of transversely isotropic bodies in which the stress components are assumed to be polynomial functions of the kinematic tensors defining the deformation and its successive time rates of variation. When an arbitrary number of kinematic tensors is involved, a preliminary reduction of the expression for the stress is achieved by using the results of classical invariant theory. The resulting expression is reduced to closed form by using the Cayley-Hamilton theorem and matrix relations derived from this by RIVLIN (1955) and by SPENCER & RIVLIN (1959a, b).

A feature of the analysis for transversely isotropic bodies is the employment of two alternative systems of invariants. In one of these, the kinematic tensors are regarded as a system of symmetric 3×3 matrices; in the other, these tensors are decomposed into a system of symmetric and unsymmetric 2×2 matrices. This latter resolution enables rapid reductions of the basic invariant system to be achieved by using comparatively simple results which depend upon the Cayley Hamilton theorem for a 2×2 matrix. Relations between the two invariant systems are readily obtained by using the symbolic methods of classical invariant theory (WEITZENBÖCK 1923, TURNBULL 1928).

In the present paper the analysis is extended to include the case where the stress matrix is a polynomial not only in the elements of a set of kinematic tensors but also in the components of a number of vector fields. Attention is here confined to the invariance problem which is thus introduced and the possible physical significance of such vectors is not examined. It is feasible, however, that they could arise as the result of electric or magnetic fields, non-uniform temperature distributions or the successive time rates of change of quantities of this type. When attention is confined to polynomial representations a distinction may be made between materials which exhibit only rotational symmetry about a given direction at each point and transversely isotropic materials in which there is also symmetry with respect to planes containing this axis. These two cases are examined separately. As in the earlier work, convected coordinate systems are employed; this ensures that the necessary conditions are satisfied for invariance under rigid body motions of the entire system (see, for example, OLDROYD 1950).

2. Mechanical properties dependent upon vectors and tensors

We consider a homogeneous body which is at rest at time $t=0$ and which subsequently undergoes a continuously varying deformation. A system of curvilinear coordinates ϑ^i is associated with elements of the material and moves with the body as it is deformed. The covariant and contravariant metric tensors for this system at the initial time $t=0$ are $\gamma_{ij}=\gamma_{ij}(\vartheta^r, 0)$ and $\gamma^{ij}=\gamma^{ij}(\vartheta^r, 0)$ respectively, the corresponding quantities at the current time t being $\Gamma_{ij}=\Gamma_{ij}(\vartheta^r, t)$ and $\Gamma^{ij}=\Gamma^{ij}(\vartheta^r, t)$. In terms of these quantities the covariant strain tensor η_{ij} and rate of strain tensors $\alpha_{ij}^{(r)}$ $(r=1, 2, \ldots, m)$ are defined by

$$\begin{gathered}\eta_{ij}=\tfrac{1}{2}(\Gamma_{ij}-\gamma_{ij}),\\ \alpha_{ij}^{(r)}=\frac{D^r\eta_{ij}}{Dt^r}=\frac{1}{2}\frac{D^r\Gamma_{ij}}{Dt^r},\\ (r=1,2,\ldots,m),\end{gathered} \tag{2.1}$$

where D/Dt denotes differentiation with respect to t holding the convected coordinates ϑ^i constant. In addition to these kinematic tensors, we suppose that the mechanical properties of the body depend also upon a number of vector fields $\boldsymbol{v}_r$ $(r=1, 2, \ldots, n)$, the covariant components of the vector $\boldsymbol{v}_r$ referred to the system ϑ^i being $v_i^{(r)}$. These vector fields may arise as the result of magnetization, electric polarization or other physical phenomena; some may be introduced as the gradients of scalar fields such as temperature while others may enter as the time rates of change of physical vectors. In the latter case such time derivatives would be formed with the operator D/Dt, the process being analogous to that defined by (2.1). We are not here concerned with the exact nature of the fields which can occur or their significance; we examine only the invariance problem which arises from the assumption that they are present in the expression for the stress tensor. The contravariant components of this tensor, referred to the system ϑ^i, are denoted by τ^{ij}.

In considering properties of the undeformed body it is convenient to form contravariant, mixed and covariant components of the kinematic and mechanical tensors and vectors already defined by raising and lowering indices with the initial metric tensors γ^{ij}, γ_{ij}. When the coordinates ϑ^i coincide with a rectangular Cartesian system $x^i(=x_i)$ in the undeformed body at time $t=0$, $\gamma_{ij}=\gamma^{ij}=\delta_{ij}$ and the corresponding components of any given tensor are equal. We denote by e_{ij}, $a_{ij}^{(r)}$, t^{ij} the components of the tensors η_{ij}, $\alpha_{ij}^{(r)}$, τ^{ij} respectively, and by $v_i^{(r)}$ those of the vector $\boldsymbol{v}_r$ when $\vartheta^i=x_i$. We define a homogeneous, uniformly aeolotropic material by the tensor relation

$$t^{ij}=f^{ij}(e_{rs}, a_{pq}^{(1)}, \ldots, a_{uv}^{(m)}, v_k^{(1)}, \ldots, v_l^{(n)}) \qquad (f^{ij}=f^{ji}) \tag{2.2}$$

where the functions f^{ij} are polynomials in the arguments indicated. Any coefficients occurring in these polynomials are material constants which are independent of position throughout the material and do not depend upon the deformation. This proviso, together with the fact that all kinematic and mechanical tensors occurring in (2.2) are referred to the initial Cartesian system x_i, is sufficient to ensure that the material is uniformly aeolotropic. The use of convected coordinates ensures that the relations (2.2) are invariant under all rigid body

motions of the entire mechanical system. This aspect has been examined by OLDROYD (1950), RIVLIN & ERICKSEN (1955), GREEN & RIVLIN (1957) and others.

If ϑ^i is a convected system not coincident initially with the Cartesian system x_i, we obtain by a tensor transformation

$$\tau^{ij} = \frac{\partial \vartheta^i}{\partial x^r} \frac{\partial \vartheta^j}{\partial x^s} t^{rs} = C^{ij}_{rs} t^{rs}, \tag{2.3}$$

where

$$C^{ij}_{rs} = \frac{1}{2}\left(\frac{\partial \vartheta^i}{\partial x^r} \frac{\partial \vartheta^j}{\partial x^s} + \frac{\partial \vartheta^i}{\partial x^s} \frac{\partial \vartheta^j}{\partial x^r}\right). \tag{2.4}$$

From the components e_{ij}, $a^{(r)}_{ij}$ of the kinematic matrices we form the symmetric matrices

$$\boldsymbol{E} = \|e_{ij}\|, \quad \boldsymbol{A}_r = \|a^{(r)}_{ij}\|, \tag{2.5}$$

and for a given pair of values of i, j the transformation matrix

$$\boldsymbol{C} = \|C_{rs}\| \qquad (C_{rs} = C^{ij}_{rs}). \tag{2.6}$$

Symmetry properties of the material are investigated by examining the behaviour of the relations (2.2) under rotations and reflections of the x_i-axes in the undeformed body. Such transformations are defined by subgroups of the full orthogonal group

$$\bar{x}_i = M_{ij} x_j, \quad M_{ik} M_{jk} = M_{ki} M_{kj} = \delta_{ij}, \tag{2.7}$$

and under all such transformations the matrices $\boldsymbol{E}$, $\boldsymbol{A}_r$, $\boldsymbol{C}$ behave similarly. We may therefore write

$$\bar{\chi}_{ij} = M_{ir} M_{js} \chi_{rs}, \tag{2.8}$$

where χ_{rs} may represent elements of any one of the matrices (2.5) or (2.6). Similarly for the vectors $\boldsymbol{v}_r$ we have

$$\bar{v}^{(r)}_i = M_{ik} v^{(r)}_k. \tag{2.9}$$

3. Transversely isotropic bodies

As in the previous paper (ADKINS 1960), general forms of the stress deformation relations for transversely isotropic bodies are deduced from results in classical invariant theory. We consider the behaviour of the system of linear and quadratic forms

$$\begin{aligned} &F_1 = A_i x_i, \qquad F_2 = B_i x_i, \ldots, F_n = N_i x_i, \\ &f_1 = a_{ij} x_i x_j, \quad f_2 = b_{ij} x_i x_j, \ldots, f_m = m_{ij} x_i x_j \qquad (i, j = 1, 2, 3) \end{aligned} \tag{3.1}$$

under the group of orthogonal transformations

$$\bar{x}_1 = x_1, \quad \bar{x}_\varrho = M_{\varrho\sigma} x_\sigma, \tag{3.2}$$

where

$$M_{\varrho\sigma} M_{\varrho\lambda} = M_{\sigma\varrho} M_{\lambda\varrho} = \delta_{\sigma\lambda}, \quad |M_{\sigma\varrho}| = +1, \tag{3.3}$$

and here and subsequently greek indices are restricted to take the values 2, 3 while latin indices cover the full range 1, 2, 3. Under the transformations (3.2)

the forms (3.1) become

$$F_1 = \bar{A}_i \bar{x}_i, \qquad F_2 = \bar{B}_i \bar{x}_i, \ldots, F_n = \bar{N}_i \bar{x}_i, \tag{3.4}$$

$$f_1 = \bar{a}_{ij} \bar{x}_i \bar{x}_j, \qquad f_2 = \bar{b}_{ij} \bar{x}_i \bar{x}_j, \ldots, f_m = \bar{m}_{ij} \bar{x}_i \bar{x}_j, \tag{3.5}$$

where

$$\begin{gathered} \bar{A}_1 = A_1, \qquad \bar{A}_\varrho = M_{\varrho\sigma} A_\sigma, \\ \bar{a}_{11} = a_{11}, \qquad \bar{a}_{1\lambda} = M_{\lambda\mu} a_{1\mu}, \qquad \bar{a}_{\lambda\varrho} = M_{\lambda\mu} M_{\varrho\sigma} a_{\mu\sigma}, \end{gathered} \tag{3.6}$$

and we have similar results for the coefficients in

$$F_2, \ldots, F_n, \qquad f_2, \ldots, f_m.$$

For the examination of invariance properties, it is convenient to use the symbolic representation of classical invariant theory for the coefficients of the quadratic forms. We therefore write

$$a_{ij} = a_{ji} = a_i a_j = a'_i a'_j = a''_i a''_j = \cdots, \tag{3.7}$$

with corresponding expressions for $b_{ij}, \ldots, m_{ij}$. The symbols a_i, a'_i, $a''_i, \ldots, m_i$, m'_i, $m''_i, \ldots$ and the coefficients $A_i, B_i, \ldots, N_i$ behave similarly under corresponding transformations of the orthogonal group. In particular, for the subgroup defined by (3.2), (3.3) we have

$$\bar{a}_1 = a_1, \qquad \bar{a}_\varrho = M_{\varrho\sigma} a_\sigma, \tag{3.8}$$

corresponding to the first of (3.6). Furthermore, any function of the coefficients $a_{ij}, b_{ij}, \ldots, m_{ij}$ may be represented in terms of the corresponding symbols and this symbolic representation gives a unique determination provided different symbols of a given set are used to represent a given coefficient each time it occurs. For example

$$a_{ij} a_{ij} = a_i a'_i a_j a'_j, \qquad a_{ii} a_{jj} = a_i a_i a'_j a'_j.$$

To the system of forms (3.4), (3.5) we adjoin the linear form

$$L \equiv l_i x_i = x_1, \tag{3.9}$$

where

$$(l_1, l_2, l_3) = (1, 0, 0). \tag{3.10}$$

It then follows from results given by Weitzenböck (1923) and others that any polynomial in the quantities $A_i, B_i, \ldots, N_i, a_{ij}, b_{ij}, \ldots, m_{ij}$ which is invariant under all transformations of the group (3.2) can be represented symbolically as a sum of products either of the quantities

$$\begin{gathered} (\xi\eta\zeta) = \begin{vmatrix} \xi_1 & \xi_2 & \xi_3 \\ \eta_1 & \eta_2 & \eta_3 \\ \zeta_1 & \zeta_2 & \zeta_3 \end{vmatrix}, \qquad (\xi\eta l), \qquad \text{(i)} \\ (\xi\eta) = \xi_i \eta_i, \qquad (\xi l) = \xi_i l_i = \xi_1, \qquad \text{(ii)} \end{gathered} \tag{3.11}$$

or of the expressions

$$(\xi\eta l), \qquad (\xi|\eta) = \xi_\alpha \eta_\alpha, \qquad (\xi l) = \xi_1. \tag{3.12}$$

In (3.11), (3.12) ξ_i, η_i, ζ_i may represent any of the symbols a_i, a'_i, $a''_i, \ldots, m_i$, m'_i, $m''_i, \ldots$ or the coefficients $A_i, B_i, \ldots, N_i$. Relations between invariants formed

from the systems (3.11), (3.12) may be obtained by writing

$$\begin{aligned}(\xi\eta) &= (\xi|\eta) + (\xi l)(\eta l), \\ (\xi\eta\zeta) &= (\xi\eta l)(l\zeta) - (\xi\zeta l)(l\eta) + (\eta\zeta l)(l\xi).\end{aligned} \tag{3.13}$$

If we consider polynomials in the coefficients of (3.1) which are invariant not only under transformations of the group (3.2) but also under the reflection

$$(\bar{x}_1, \bar{x}_2, \bar{x}_3) = (x_1, x_2, -x_3), \tag{3.14}$$

the determinants $(\xi\eta\zeta)$, $(\xi\eta l)$ must be omitted from the symbolic representations. In place of (3.11), (3.12) we then have

$$(\xi\eta) = \xi_i\eta_i, \qquad (l\xi) = \xi_1, \tag{3.15}$$

$$(\xi|\eta) = \xi_\alpha\eta_\alpha, \qquad (l\xi) = \xi_1, \tag{3.16}$$

respectively.

From the definitions of § 2 it follows that the coefficients a_{ij} and the elements of the matrices $\boldsymbol{C}, \boldsymbol{E}, \boldsymbol{A}_r$ behave similarly under all transformation of the orthogonal group; the same is true of the coefficients A_i and the vector components $v_i^{(r)}$. Furthermore, the transformations (3.2) and (3.14) may be used to describe a rotation of the x_i-axes of § 2 about the x_1-axis and a reflection in the x_1, x_2-plane respectively. It follows that if from the symbolic quantities (3.11) or (3.12) we construct invariants of the coefficients in (3.1), and in the resulting expressions replace $a_{ij}, b_{ij}, \ldots, m_{ij}$ by elements of $\boldsymbol{C}, \boldsymbol{E}, \boldsymbol{A}_r$ and $A_i, B_i, \ldots, N_i$ by elements of $\boldsymbol{v}_r$ we obtain a system of invariants and coefficients appropriate to materials possessing the rotational symmetry described by (3.2). A similar procedure based on (3.15) or (3.16) yields the corresponding system for bodies transversely isotropic relative to the x_1-axis, these being characterised by the transformations (3.2), (3.14).

Since the product of two symbolic determinants $(\xi\eta\zeta)$ $(\xi'\eta'\zeta')$ can be expressed as a polynomial in the scalar products (ii) of (3.11) it follows that any polynomial invariant of the rotational symmetry group formed from the quantities (3.11) can be represented symbolically as

$$(\xi\eta\zeta) J_1 + (\xi\eta l) J_2 + J_3,$$

where J_1, J_2, J_3 are polynomials in the scalar products $(\xi\eta)$, (ξl). Similarly for any polynomial invariant formed from the quantities (3.12) we may obtain a symbolic representation

$$(\xi\eta l) J_1' + J_2',$$

where J_1', J_2' are polynomials in the products $(\xi|\eta)$, (ξl) alone. Evidently, if the polynomial basis is abandoned, the factors $(\xi\eta\zeta)$, $(\xi\eta l)$ may be omitted and the invariants for the rotation group may be constructed from the quantities (3.15) or (3.16).

4. Some relations for matrices

For subsequent reductions we require the following results in the theory of matrix algebra which are a consequence of the Cayley-Hamilton theorem and are based on analysis given by Rivlin (1955):

(i) Any matrix polynomial $\boldsymbol{P}$ in n 2×2 matrices $\boldsymbol{B}_k$ $(k=1, 2, \ldots, n)$ can be expressed in the form

$$\boldsymbol{P} = \varphi_0 \boldsymbol{I} + \sum_{j=1}^{n} \varphi_j \boldsymbol{B}_j + \sum_{j=1}^{n} \sum_{k=1}^{n} \varphi_{jk} \boldsymbol{B}_j \boldsymbol{B}_k, \tag{4.1}$$

where φ_0, φ_j and φ_{jk} are scalars expressible as polynomials in the elements of $\boldsymbol{B}_i$ and $\boldsymbol{I}$ is the unit matrix. If $\boldsymbol{P}$ is homogeneous and of degree m in the elements of the matrices $\boldsymbol{B}_i$ then each term of the right hand member is also homogeneous and of degree m in these elements. Furthermore, from the method of proof given by RIVLIN, it follows that if $\boldsymbol{P}$ is a matrix product of the form $\boldsymbol{B}_i^{\alpha_i} \boldsymbol{B}_j^{\alpha_j} \ldots \boldsymbol{B}_k^{\alpha_k}$ where $\alpha_i, \alpha_j \ldots \alpha_k$ are positive integers, then $\varphi_0, \varphi_j, \varphi_{jk}$ are all polynomials in traces of products and powers of the matrices $\boldsymbol{B}_k$.

(ii) Any invariant I of the form $\operatorname{tr} \boldsymbol{B}_i \boldsymbol{B}_j \ldots \boldsymbol{B}_k$ formed from the system of 2×2 matrices $\boldsymbol{B}_i$ is expressible as a polynomial in

$$\operatorname{tr} \boldsymbol{B}_i, \quad \operatorname{tr} \boldsymbol{B}_i \boldsymbol{B}_j, \quad \operatorname{tr} \boldsymbol{B}_i \boldsymbol{B}_j \boldsymbol{B}_k.$$

This follows immediately from (i) by forming $\operatorname{tr} \boldsymbol{B}_i \boldsymbol{P}$ and writing $\boldsymbol{P} = \boldsymbol{B}_j \ldots \boldsymbol{B}_k$. This gives

$$I = \varphi_0 \operatorname{tr} \boldsymbol{B}_i + \sum_{j=1}^{n} \varphi_j \operatorname{tr} \boldsymbol{B}_i \boldsymbol{B}_j + \sum_{j=1}^{n} \sum_{k=1}^{n} \varphi_{jk} \operatorname{tr} \boldsymbol{B}_i \boldsymbol{B}_j \boldsymbol{B}_k,$$

and if I is of degree $m(>3)$ in the elements of $\boldsymbol{B}_1, \boldsymbol{B}_2, \ldots, \boldsymbol{B}_n$ the invariant functions $\varphi_0, \varphi_j, \varphi_{jk}$ are of degrees less than m. Repetition of this procedure gives the result stated.

(iii) Any polynomial invariant formed from the elements of a system of 2×2 *symmetric* matrices $\boldsymbol{B}_k$ $(k=1, 2, \ldots, n)$ can be expressed as a polynomial in

$$\operatorname{tr} \boldsymbol{B}_i, \quad \operatorname{tr} \boldsymbol{B}_i \boldsymbol{B}_k.$$

(iv) Any polynomial invariant formed from the elements of a system of 2×2 symmetric matrices $\boldsymbol{B}_k$ which involves also the elements of an unsymmetric matrix $\boldsymbol{B}$ linearly can be expressed as a polynomial in

$$\operatorname{tr} \boldsymbol{B}_i, \quad \operatorname{tr} \boldsymbol{B}_i \boldsymbol{B}_k, \quad \operatorname{tr} \boldsymbol{B}, \quad \operatorname{tr} \boldsymbol{B}_i \boldsymbol{B}, \quad \operatorname{tr} \boldsymbol{B}_i \boldsymbol{B}_k \boldsymbol{B}.$$

This result follows as a corollary from (ii) and (iii).

(v) If $\boldsymbol{A}$ and $\boldsymbol{B}$ are two 3×3 matrices then

$$\boldsymbol{A} \boldsymbol{B} \boldsymbol{A} = -\boldsymbol{A}^2 \boldsymbol{B} - \boldsymbol{B} \boldsymbol{A}^2 + \boldsymbol{P}(2), \tag{4.2}$$

where $\boldsymbol{P}(2)$ is a matrix polynomial of degree two in $\boldsymbol{A}$, $\boldsymbol{B}$, the coefficients in this polynomial being scalar polynomial invariants formed from the elements of $\boldsymbol{A}$ and $\boldsymbol{B}$.

The results (i) to (iv) depend upon the Cayley-Hamilton theorem

$$\boldsymbol{A}^2 = \boldsymbol{A} \operatorname{tr} \boldsymbol{A} + \tfrac{1}{2} \boldsymbol{I} [\operatorname{tr} \boldsymbol{A}^2 - (\operatorname{tr} \boldsymbol{A})^2] \tag{4.3}$$

for a 2×2 matrix $\boldsymbol{A}$; (v) depends upon the corresponding result for a 3×3 matrix.

5. Reduction of invariants for transversely isotropic bodies

From the symbolic products (3.15) we may construct invariants of the following types:

$$\begin{array}{llll} a_{ij}\, b_{jk} \ldots c_{lm}\, f_{mi}, & \text{(i)} & a_{1i}\, b_{ik} \ldots c_{lm}\, f_{m1}, & \text{(ii)} \\ A_i\, a_{ij}\, b_{jk} \ldots c_{lm}\, f_{m1}, & \text{(iii)} & A_i\, a_{ij}\, b_{jk} \ldots c_{lm}\, f_{mn}\, B_n, & \text{(iv)} \end{array} \tag{5.1}$$

Type (i) arises from products of the factors $(\xi\,\eta)$ alone in which each symbol occurs twice; these symbols therefore combine, as in (3.7), to yield coefficients $a_{ij}, b_{ij}, \ldots, m_{ij}$ of the quadratic forms in (3.1). Type (iv) also arises from the products $(\xi\,\eta)$ but two of the symbols in this product occur only once giving two coefficients of the linear forms in (3.1). In the symbolic forms representing (ii) there are two factors $(\xi\, l)$ containing l, and each symbol $\xi, \eta, \ldots$ occurs twice; in (iii) there is a single factor $(\xi\, l)$ and one of the symbols $\xi, \eta, \ldots$ is not repeated. The coefficients $A_1, B_1, \ldots, N_1, a_{11}, b_{11}, \ldots, m_{11}$ are also invariant. It is readily seen that products containing more than two factors $(\xi\, l)$ can be written as products of simpler invariants of the types (5.1); a similar remark applies to products in which more than two of the symbols $\xi, \eta, \ldots$ occur only once. From the symbolic products (3.16) we may construct the invariants

$$\begin{array}{llll} a_{\alpha\beta}\, b_{\beta\gamma} \ldots c_{\varrho\lambda}\, f_{\lambda\alpha}, & \text{(i)} & a_{1\alpha}\, b_{\alpha\beta} \ldots c_{\varrho\lambda}\, f_{\lambda 1}, & \text{(ii)} \\ A_\alpha\, a_{\alpha\beta}\, b_{\beta\gamma} \ldots c_{\varrho\lambda}\, f_{\lambda 1}, & \text{(iii)} & A_\alpha\, a_{\alpha\beta}\, b_{\beta\gamma} \ldots c_{\varrho\lambda}\, f_{\lambda\mu}\, B_\mu, & \text{(iv)} \end{array} \tag{5.2}$$

analogous to (5.1).

The invariant system (5.1) may be reduced further by making use of the Cayley-Hamilton theorem for a 3×3 matrix and the generalizations derived by RIVLIN (1955) and by SPENCER & RIVLIN (1959a, b, 1960). Reductions of the system (5.2) are based on the Cayley-Hamilton theorem for a 2×2 matrix. Expressions for the invariants (5.1) in terms of (5.2) are obtained by converting each of the coefficients into symbolic notation and expanding the expressions obtained with the help of (3.13); the converse procedure yields expressions for the invariants (5.2) in terms of (5.1). Relations of this kind have been derived in an earlier paper (ADKINS 1960) and a general method of reduction of the invariants (i) and (ii) of (5.1) and (5.2) has there been indicated. We are here concerned mainly with the additional invariants arising from types (iii) and (iv) and consider first the forms (5.2).

We allow $a_{ij}, b_{ij}, \ldots, A_i, B_i, \ldots$ to denote typical coefficients of the forms (3.1), it being understood that different letters denote coefficients of different forms, no two forms being identical. From these coefficients we construct a system of symmetric 2×2 matrices

$$\begin{aligned} &\boldsymbol{b}_i = \|a_{\alpha\beta}\|, \quad \boldsymbol{c}_i = \|a_{1\alpha}\, a_{\beta 1}\| \qquad (i = 1, 2, \ldots, m), \\ &\boldsymbol{d}_i = \|A_\alpha\, A_\beta\| \qquad (i = 1, 2, \ldots, n), \end{aligned} \tag{5.3}$$

and a system of unsymmetric 2×2 matrices

$$\begin{aligned} &\boldsymbol{b}_{ik} = \|a_{1\alpha}\, b_{\beta 1}\| \qquad (i, k = 1, 2, \ldots, m;\; i < k;\; a \neq b), \\ &\boldsymbol{c}_{ik} = \|A_\alpha\, a_{\beta 1}\| \qquad (i = 1, 2, \ldots, n;\; k = 1, 2, \ldots, m), \\ &\boldsymbol{d}_{ik} = \|A_\alpha\, B_\beta\| \qquad (i, k = 1, 2, \ldots, n;\; i < k;\; A \neq B). \end{aligned} \tag{5.4}$$

We thus obtain m, m and n matrices of types $\boldsymbol{b}_i$, $\boldsymbol{c}_i$, $\boldsymbol{d}_i$ respectively and $\frac{1}{2}m(m-1)$, $m\,n$ and $\frac{1}{2}n(n-1)$ of types $\boldsymbol{b}_{ik}$, $\boldsymbol{c}_{ik}$ and $\boldsymbol{d}_{ik}$ respectively. It is sufficient for the subsequent analysis to omit the matrices $\|b_{1\alpha}\,a_{\beta 1}\|$, $\|B_\alpha\,A_\beta\|$ which are the transposes respectively of $\boldsymbol{b}_{ik}$, $\boldsymbol{d}_{ik}$ in (5.4).

From (5.2) we observe that any polynomial invariant formed from the coefficients of (3.1) is expressible as a polynomial in the traces of products of the matrices (5.3), (5.4) which either involve the set $\boldsymbol{b}_i$ alone or at most one of the matrices $\boldsymbol{c}_i$, $\boldsymbol{d}_i$, $\boldsymbol{b}_{ik}$, $\boldsymbol{c}_{ik}$, $\boldsymbol{d}_{ik}$ linearly. It follows at once from the results (iii) and (iv) of § 4 that each of the invariants (5.2) may be expressed as a polynomial in*

$$\left.\begin{array}{lll}
a_{11}, \quad a_{\alpha\alpha}, \quad a_{\alpha\beta}\,b_{\beta\alpha}, \quad a_{1\alpha}\,b_{\alpha 1}, & & \text{(i)}\\
a_{1\alpha}\,b_{\alpha\beta}\,c_{\beta 1}, & & \text{(ii)}\\
a_{1\alpha}\,a_{\alpha\beta}\,b_{\beta\gamma}\,c_{\gamma 1} & (a \neq b,\ a \neq c), & \text{(iii)}\\
a_{1\alpha}\,b_{\alpha\beta}\,c_{\beta\gamma}\,d_{\gamma 1} & (a, b, c, d \text{ all different}), & \text{(iv)}\\
A_1, \quad A_\alpha\,a_{\alpha 1}, \quad A_\alpha\,a_{\alpha\beta}\,b_{\beta 1}, & & \text{(v)}\\
A_\alpha\,a_{\alpha\beta}\,b_{\beta\gamma}\,c_{\gamma 1} & (a \neq b,\ a \neq c), & \text{(vi)}\\
A_\alpha\,B_\alpha, \quad A_\alpha\,B_\beta\,a_{\alpha\beta}, & & \text{(vii)}\\
A_\alpha\,B_\beta\,a_{\beta\gamma}\,b_{\gamma\alpha} & (A \neq B,\ a \neq b). & \text{(viii)}
\end{array}\right\} \tag{5.5}$$

The restriction $a \neq b$ in (iii), $b \neq c$ in (iv) and $a \neq b$ in (vi) and (viii) follow as an immediate consequence of the Cayley-Hamilton theorem (4.3). To obtain the remaining restrictions we need to use (4.2). For example, if we express the invariant (vi) in terms of those of the system (5.1) by converting into symbolic notation and using (3.13) we obtain

$$\begin{aligned}
A_\alpha\,a_{\alpha\beta}\,b_{\beta\gamma}\,c_{\gamma 1} &= (A|a)\,(a|b)\,(b|c)\,(c\,l)\\
&= [(A\,a) - (A\,l)\,(a\,l)]\,[(a\,b) - (a\,l)\,(b\,l)]\,[(b\,c) - (b\,l)\,(c\,l)]\,(c\,l)\\
&= A_i\,a_{ij}\,b_{jk}\,c_{k1} - A_1\,a_{1i}\,b_{ij}\,c_{j1} - A_i\,a_{i1}\,b_{1j}\,c_{j1} -\\
&\quad - A_i\,a_{ij}\,b_{j1}\,c_{11} + A_1\,a_{11}\,b_{1i}\,c_{i1} +\\
&\quad + A_1\,a_{1i}\,b_{i1}\,c_{11} + A_i\,a_{i1}\,b_{11}\,c_{11} - A_1\,a_{11}\,b_{11}\,c_{11}.
\end{aligned} \tag{5.6}$$

If $a_{ij} = c_{ij}$ the first term in the expansion (5.6) may be reduced by means of (4.2) giving

$$A_i\,a_{ij}\,b_{jk}\,a_{k1} = -A_i\,a_{ij}\,a_{jk}\,b_{k1} - A_i\,b_{ij}\,a_{jk}\,a_{k1} + \Psi, \tag{5.7}$$

where Ψ is a polynomial formed from invariants of the second degree or lower in a_{ij}, b_{ij}, these invariants being irreducible. With the help of the symbolic

---

* Since this paper was prepared, a paper has appeared in which A. C. PIPKIN R. S. RIVLIN [Arch. Rational Mech. Anal. **4**, 129 (1959)] give the invariant system for a *single* matrix and a system of vectors. They treat the problem from a slightly different view point from that of the present paper, and their invariant system may be regarded as a special case of that given above.

method, the first term on the right-hand side of (5.7) may be written

$$A_i\, a_{ij}\, a_{jk}\, b_{k1} = A_\alpha\, a_{\alpha\beta}\, a_{\beta\gamma}\, b_{\gamma 1} + \Psi' = \Psi'', \tag{5.8}$$

where Ψ', Ψ'' are polynomials in invariants of the second degree or less in the elements a_{ij}, b_{ij}, the final result following as a result of the Cayley-Hamilton theorem (4.3). Application of the symbolic process to the term $A_i\, b_{ij}\, a_{jk}\, a_{k1}$ in (5.7) yields an expression for this quantity in terms of the invariant system (5.5) thereby completing the required reduction. A corresponding procedure applied to the invariants (iv) of (5.5) leads to the restrictions indicated upon the letters a, b, c, d (ADKINS 1960).

From the system (5.5) we may, by using symbolic methods, construct an equivalent system of invariants of the types (5.1) in which the repeated suffices cover the full range 1, 2, 3. We thus derive the alternative scheme

$$\left.\begin{array}{lll}
a_{ii}, \quad a_{ij}\, a_{ji}, \quad a_{ij}\, a_{jk}\, a_{ki}, & & \text{(i)}\\
a_{ij}\, b_{ji}, \quad a_{ij}\, a_{jk}\, b_{ki}, \quad a_{ij}\, a_{jk}\, b_{kl}\, b_{li} & (a \neq b), & \text{(ii)}\\
a_{11}, \quad a_{1i}\, a_{i1}, \quad a_{1i}\, b_{i1}, \quad a_{1i}\, a_{ij}\, b_{j1} & (a \neq b), & \text{(iii)}\\
a_{1i}\, b_{ik}\, c_{k1}, \quad a_{11}\, b_{ij}\, c_{jk}\, c_{k1}, \quad a_{1i}\, b_{ij}\, c_{jk}\, d_{k1} & (a, b, c, d \text{ all different}), & \text{(iv)}\\
A_1, \quad A_i\, a_{i1}, \quad A_i\, a_{ij}\, b_{j1}, & & \text{(v)}\\
A_i\, a_{ij}\, b_{jk}\, c_{k1} \quad (a \neq b,\ a \neq c), & & \text{(vi)}\\
A_i\, B_i, \quad A_i\, B_j\, a_{ij}, & & \text{(vii)}\\
A_i\, B_j\, a_{jk}\, b_{ki} \quad (A \neq B,\ a \neq b), & & \text{(viii)}
\end{array}\right\} \tag{5.9}$$

in which (v) to (viii) follow directly from the corresponding invariants of (5.5) and (i) to (iv) have been given in the paper previously cited.

In applying these results to restrict the form of the stress deformation relations, we first observe that under transformations of the x_i-axes, each of the stress components τ^{ij} may be regarded as a scalar invariant formed from the matrices $\boldsymbol{C}, \boldsymbol{E}(=\boldsymbol{A}_0)$, A_r $(r=1, 2, \ldots, n)$ and the vectors $\boldsymbol{v}_s$ $(s=1, 2, \ldots, m)$ which is form invariant under the transformations (3.2), (3.14) and involves the elements of $\boldsymbol{C}$ linearly. We may therefore write

$$\tau^{ij} = \sum_{r=1}^{M} P_{(r)}^{ij}\, \varphi_r, \tag{5.10}$$

$$P_{(r)}^{ij} = P_{(r)}^{ij}(\boldsymbol{C}, \boldsymbol{E}, \boldsymbol{A}_p, \boldsymbol{v}_q), \qquad \varphi_r = \varphi_r(\boldsymbol{E}, \boldsymbol{A}_p, \boldsymbol{v}_q), \tag{5.11}$$

where $P_{(r)}^{ij}$, φ_r are polynomial invariant functions of the arguments indicated and $P_{(r)}^{ij}$ are linear and homogeneous in the elements of $\boldsymbol{C}$ (i.e. C_{rs}^{ij}). It follows that the functions φ_r are polynomials in the invariants formed from (5.9) by replacing the coefficients a_{ij}, b_{ij}, c_{ij}, d_{ij} in every permissible manner by the corresponding elements of $\boldsymbol{E}, \boldsymbol{A}_r$ and the coefficients A_i, B_i similarly by elements of the vectors $\boldsymbol{v}_r$.

The polynomials φ_r thus become functions of the systems of invariants

$$\begin{array}{llll}
\operatorname{tr}\boldsymbol{A}_i, \quad \operatorname{tr}\boldsymbol{A}_i^2, \quad \operatorname{tr}\boldsymbol{A}_i^3, & & \text{(i)} & \\
\operatorname{tr}\boldsymbol{A}_i\boldsymbol{A}_j, \quad \operatorname{tr}\boldsymbol{A}_i^2\boldsymbol{A}_j, \quad \operatorname{tr}\boldsymbol{A}_i\boldsymbol{A}_j^2, \quad \operatorname{tr}\boldsymbol{A}_i^2\boldsymbol{A}_j^2, & & \text{(ii)} & \\
[\boldsymbol{A}_i]_{11}, \quad [\boldsymbol{A}_i^2]_{11}, & & \text{(iii)} & \\
[\boldsymbol{A}_i\boldsymbol{A}_j]_{11}, \quad [\boldsymbol{A}_i^2\boldsymbol{A}_j]_{11}, \quad [\boldsymbol{A}_i\boldsymbol{A}_j^2]_{11} & (i \neq j), & \text{(iv)} & \\
[\boldsymbol{A}_i\boldsymbol{A}_j\boldsymbol{A}_k]_{11}, \quad [\boldsymbol{A}_j\boldsymbol{A}_i\boldsymbol{A}_k]_{11}, \quad [\boldsymbol{A}_i\boldsymbol{A}_j\boldsymbol{A}_k^2]_{11}, & & & (5.12)^{*} \\
[\boldsymbol{A}_i\boldsymbol{A}_k\boldsymbol{A}_j^2]_{11}, \quad [\boldsymbol{A}_j\boldsymbol{A}_k\boldsymbol{A}_i^2]_{11} & (i, j, k \text{ all different}), & \text{(v)} & \\
[\boldsymbol{A}_i\boldsymbol{A}_j\boldsymbol{A}_k\boldsymbol{A}_l]_{11}, \quad [\boldsymbol{A}_i\boldsymbol{A}_j\boldsymbol{A}_l\boldsymbol{A}_k]_{11}, \quad [\boldsymbol{A}_i\boldsymbol{A}_k\boldsymbol{A}_j\boldsymbol{A}_l]_{11}, & & & \\
[\boldsymbol{A}_i\boldsymbol{A}_k\boldsymbol{A}_l\boldsymbol{A}_j]_{11}, \quad [\boldsymbol{A}_i\boldsymbol{A}_l\boldsymbol{A}_j\boldsymbol{A}_k]_{11}, \quad [\boldsymbol{A}_j\boldsymbol{A}_i\boldsymbol{A}_k\boldsymbol{A}_l]_{11} & & & \\
 & (i, j, k, l \text{ all different}), & \text{(vi)} &
\end{array}$$

$$\begin{array}{llll}
[\boldsymbol{v}_r]_1, \quad [\boldsymbol{v}_r\boldsymbol{A}_i]_1, \quad [\boldsymbol{v}_r\boldsymbol{A}_i^2]_1, & & \text{(i)} & \\
[\boldsymbol{v}_r\boldsymbol{A}_i\boldsymbol{A}_j]_1, \quad [\boldsymbol{v}_r\boldsymbol{A}_i\boldsymbol{A}_j^2]_1, \quad [\boldsymbol{v}_r\boldsymbol{A}_i\boldsymbol{A}_j\boldsymbol{A}_k]_1, & & \text{(ii)} & \\
\operatorname{tr}\boldsymbol{v}_r^2, \quad \operatorname{tr}\boldsymbol{v}_r^2\boldsymbol{A}_i, & & \text{(iii)} & (5.13) \\
\operatorname{tr}\boldsymbol{v}_r\boldsymbol{v}_s, \quad \operatorname{tr}\boldsymbol{v}_r\boldsymbol{v}_s\boldsymbol{A}_i, \quad \operatorname{tr}\boldsymbol{v}_r\boldsymbol{v}_s\boldsymbol{A}_i\boldsymbol{A}_j, & & & \\
 & (r \neq s;\ i, j, k \text{ all different}), & \text{(iv)} &
\end{array}$$

obtained by taking all possible choices of one, two, three or four as appropriate, of the letters i, j, k, l from the numbers 0 to m ($\boldsymbol{A}_0 = \boldsymbol{E}$) and of r, s from the numbers 1 to n, together with the invariants derived from (5.13) by all possible permutations of i, j, k.

In (5.12) the symbol $[\ ]_{11}$ denotes the leading term of a matrix product corresponding to (iii) and (iv) of (5.9), in (5.13) $[\ ]_1$ indicates a product of types (v) and (vi) containing a single vector and with one suffix taking the value unity, while the last five terms of (5.13) are scalar products containing a pair of vectors. The meaning of the symbol tr has been extended to cover these terms, which correspond to (vii) and (viii) of (5.9), where summation is carried out over all suffixes.

The coefficients $P_{(r)}^{ij}$ are derived by forming from (5.9) the invariants in $\boldsymbol{C}$, $\boldsymbol{E}$, $\boldsymbol{A}_p$, $\boldsymbol{v}_q$ which involve $\boldsymbol{C}$ linearly. This yields the system

$$\begin{array}{llll}
\operatorname{tr}\boldsymbol{C}, \quad \operatorname{tr}\boldsymbol{C}\boldsymbol{A}_i, \quad \operatorname{tr}\boldsymbol{C}\boldsymbol{A}_i^2, & & \text{(i)} & \\
[\boldsymbol{C}]_{11}, \quad [\boldsymbol{C}\boldsymbol{A}_i]_{11}, \quad [\boldsymbol{C}\boldsymbol{A}_i^2]_{11}, & & \text{(ii)} & \\
[\boldsymbol{C}\boldsymbol{A}_i\boldsymbol{A}_j]_{11}, \quad [\boldsymbol{C}\boldsymbol{A}_j\boldsymbol{A}_i]_{11}, \quad [\boldsymbol{C}\boldsymbol{A}_i\boldsymbol{A}_j^2]_{11}, \quad [\boldsymbol{C}\boldsymbol{A}_j\boldsymbol{A}_i^2]_{11} & (i \neq j), & \text{(iii)} & (5.14) \\
[\boldsymbol{C}\boldsymbol{A}_i\boldsymbol{A}_j\boldsymbol{A}_k]_{11}, \quad [\boldsymbol{C}\boldsymbol{A}_i\boldsymbol{A}_k\boldsymbol{A}_j]_{11}, \quad [\boldsymbol{C}\boldsymbol{A}_j\boldsymbol{A}_i\boldsymbol{A}_k]_{11}, & & & \\
[\boldsymbol{C}\boldsymbol{A}_j\boldsymbol{A}_k\boldsymbol{A}_i]_{11}, \quad [\boldsymbol{C}\boldsymbol{A}_k\boldsymbol{A}_i\boldsymbol{A}_j]_{11}, \quad [\boldsymbol{A}_i\boldsymbol{C}\boldsymbol{A}_j\boldsymbol{A}_k]_{11} & (i, j, k \text{ all different}), & \text{(iv)} &
\end{array}$$

* For the derivation of the restrictions on the invariants (v) and (vi) reference may be made to the earlier paper.

$$\begin{array}{ll} [\boldsymbol{v}_r\,\boldsymbol{C}]_1, \quad [\boldsymbol{v}_r\,\boldsymbol{C}\,\boldsymbol{A}_i]_1, \quad [\boldsymbol{v}_r\,\boldsymbol{A}_i\,\boldsymbol{C}]_1, \quad [\boldsymbol{v}_r\,\boldsymbol{C}\,\boldsymbol{A}_i^2]_1, & \text{(i)} \\ [\boldsymbol{v}_r\,\boldsymbol{C}\,\boldsymbol{A}_i\,\boldsymbol{A}_j]_1, \quad [\boldsymbol{v}_r\,\boldsymbol{C}\,\boldsymbol{A}_j\,\boldsymbol{A}_i]_1, \quad [\boldsymbol{v}_r\,\boldsymbol{A}_i\,\boldsymbol{C}\,\boldsymbol{A}_j]_1, & \\ [\boldsymbol{v}_r\,\boldsymbol{A}_j\,\boldsymbol{C}\,\boldsymbol{A}_i]_1, \quad [\boldsymbol{v}_r\,\boldsymbol{A}_i\,\boldsymbol{A}_j\,\boldsymbol{C}]_1 \quad (i \neq j), & \text{(ii)}^{\star} \\ \operatorname{tr}\boldsymbol{C}\,\boldsymbol{v}_r^2, \quad \operatorname{tr}\boldsymbol{C}\,\boldsymbol{v}_r\,\boldsymbol{v}_s, \quad \operatorname{tr}\boldsymbol{v}_r\,\boldsymbol{v}_s\,\boldsymbol{C}\,\boldsymbol{A}_i, \quad \operatorname{tr}\boldsymbol{v}_r\,\boldsymbol{v}_s\,\boldsymbol{A}_i\,\boldsymbol{C} \quad (r \neq s). & \text{(iii)} \end{array} \qquad (5.15)$$

The coefficients (i) in (5.14) and (iii) in (5.15) are those appropriate to isotropic bodies and may be written in more familiar forms in terms of components referred to the convected system ϑ^i. Remembering (2.4) and the definitions of § 2 we obtain

$$\operatorname{tr}\boldsymbol{C} = \gamma^{ij}, \quad \operatorname{tr}\boldsymbol{C}\,\boldsymbol{A}_r = \alpha^{(r)ij}, \quad \operatorname{tr}\boldsymbol{C}\,\boldsymbol{A}_r^2 = \alpha^{(r)}{}_k^i\,\alpha^{(r)kj},$$

$$\operatorname{tr}\boldsymbol{C}\,\boldsymbol{v}_r^2 = v^{(r)i}\,v^{(r)j}, \quad \operatorname{tr}\boldsymbol{C}\,\boldsymbol{v}_r\,\boldsymbol{v}_s = \tfrac{1}{2}\left[v^{(r)i}\,v^{(s)j} + v^{(r)j}\,v^{(s)i}\right]$$

$$\operatorname{tr}\boldsymbol{v}_r\,\boldsymbol{v}_s\,\boldsymbol{C}\,\boldsymbol{A}_k = \tfrac{1}{2}\left[v^{(r)l}\,v^{(s)i}\,\alpha^{(k)}{}_l^j + v^{(r)l}\,v^{(s)j}\,\alpha^{(k)}{}_l^i\right].$$

6. Polynomial invariants for rotational symmetry

It is of interest to examine the forms of the additional invariants which would be required for a *polynomial* representation of materials possessing only rotational symmetry. The symbolic representations of these are linear in $(\xi\,\eta\,\zeta)$ or $(\xi\,\eta\,l)$ if formed from the elements (3.11) or linear in $(\xi\,\eta\,l)$ if formed from (3.12). We investigate the consequences of this latter representation.

The additional invariants for rotational symmetry are readily constructed by noticing that

$$(\xi\,\eta\,l) = \varepsilon_{ijk}\,\xi_i\,\eta_j\,l_k = \varepsilon_{\alpha\beta 1}\,\xi_\alpha\,\eta_\beta = \varepsilon_{\alpha\beta}\,\xi_\alpha\,\eta_\beta, \tag{6.1}$$

where $\varepsilon_{ijk} = +1$ or -1 according as i, j, k is an even or odd permutation of 1, 2, 3 and is zero otherwise and $\varepsilon_{\alpha\beta} = \varepsilon_{\alpha\beta 1}$. Symbolically, (6.1) may be written

$$(\xi\,\eta\,l) = (\varepsilon|\xi)\,(\varepsilon|\eta),$$

with $\varepsilon_\alpha\,\varepsilon_\beta = \varepsilon_{\alpha\beta}$. The additional invariants may therefore be expressed in terms of symbolic products of the types (3.16). As before, we obtain the forms (5.2), but in each invariant one of the elements $a_{\alpha\beta}, b_{\alpha\beta}, \ldots$ is replaced by $\varepsilon_{\alpha\beta}$. Reductions of these invariants again follow from the results of § 4. In this way we derive the systems

$$\begin{array}{ll} \varepsilon_{\alpha\beta}\,a_{\beta\gamma}\,b_{\gamma\alpha} \quad (a \neq b), & \text{(i)} \\ a_{1\alpha}\,\varepsilon_{\alpha\beta}\,b_{\beta 1} \quad (a \neq b), \quad a_{1\alpha}\,\varepsilon_{\alpha\beta}\,b_{\beta\gamma}\,c_{\gamma 1}, & \text{(ii)} \\ a_{1\alpha}\,\varepsilon_{\alpha\beta}\,A_\beta, \quad a_{1\alpha}\,\varepsilon_{\alpha\beta}\,b_{\beta\gamma}\,A_\gamma, \quad a_{1\alpha}\,b_{\alpha\beta}\,\varepsilon_{\beta\gamma}\,A_\gamma, & \text{(iii)} \\ A_\alpha\,\varepsilon_{\alpha\beta}\,B_\beta \quad (A \neq B), \quad A_\alpha\,\varepsilon_{\alpha\beta}\,a_{\beta\gamma}\,B_\gamma, & \text{(iv)} \end{array} \qquad (6.2)$$

★ The coefficient $[\boldsymbol{v}_r\,\boldsymbol{A}_j\,\boldsymbol{A}_i\,\boldsymbol{C}]_1$ can be expressed in terms of the remaining coefficients of (5.15) and is therefore excluded. This follows from the results obtained by Rivlin (1955) that the linear combination

$$\boldsymbol{C}\,\boldsymbol{A}_i\,\boldsymbol{A}_j + \boldsymbol{C}\,\boldsymbol{A}_j\,\boldsymbol{A}_i + \boldsymbol{A}_i\,\boldsymbol{C}\,\boldsymbol{A}_j + \boldsymbol{A}_j\,\boldsymbol{C}\,\boldsymbol{A}_i + \boldsymbol{A}_i\,\boldsymbol{A}_j\,\boldsymbol{C} + \boldsymbol{A}_j\,\boldsymbol{A}_i\,\boldsymbol{C}$$

is expressible as a polynomial in matrix products of the second and lower degree. For the same reason, the six products $[\boldsymbol{v}_r\,\boldsymbol{A}_i\,\boldsymbol{A}_j\,\boldsymbol{A}_k]$,of (5.13) obtained by permuting i, j, k are also linearly related.

corresponding to types (i), (ii), (iii) and (iv) respectively of (5.2). Any invariant of higher degree than those listed in the coefficients $a_{ij}, b_{ij}, c_{ij}, \ldots$ may be reduced by means of (4.1). For example, the invariant

$$a_{1\alpha}\,\varepsilon_{\alpha\beta}\,b_{\beta\gamma}\,c_{\gamma\varrho}\,d_{\varrho 1}$$

of type (ii) may be written as

$$\varepsilon_{\alpha\beta}(b_{\beta\gamma}\,c_{\gamma\varrho}\,\eta_{\varrho\alpha}),$$

where

$$\eta_{\varrho\alpha} = d_{\varrho 1}\,a_{1\alpha},$$

and the reduction follows immediately from (4.1).

References

Adkins, J. E.: Phil. Trans. Roy. Soc. A **250**, 519 (1958).
Adkins, J. E.: Arch. Rational Mech. Anal. **4**, 193 (1960).
Green, A. E., & R. S. Rivlin: Arch. Rational Mech. Anal. **1**, 1 (1957).
Oldroyd, J. G.: Proc. Roy. Soc. A **200**, 523 (1950).
Rivlin, R. S.: J. Rational Mech. Anal. **4**, 681 (1955).
Rivlin, R. S., & J. L. Ericksen: J. Rational Mech. Anal. **4**, 323 (1955).
Spencer, A. J. M., & R. S. Rivlin: Arch. Rational Mech. Anal. **2**, 309 (1959a).
Spencer, A. J. M., & R. S. Rivlin: Arch. Rational Mech. Anal. **2**, 435 (1959b).
Spencer, A. J. M., & R. S. Rivlin: Arch. Rational Mech. Anal. **4**, 214 (1960).
Turnbull, H. W.: The theory of determinants, matrices and invariants. London: Blackie & Sons, Ltd. 1928.
Weitzenböck, R.: Invarianten-Theorie. Groningen 1923.

Department of Theoretical Mechanics
University of Nottingham
England

(Received January 23, 1960)

EDITORIAL BOARD

CONTENTS

Druck der Universitätsdruckerei H. Stürtz AG., Würzburg
Printed in Germany